

Desai, A.  
101803541 Page 1  
Seq. ID 2

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 11:30:48 ; Search time 165 Seconds  
(without alignments)  
35.160 Million cell updates/sec

Title: US-10-803-541-2

Perfect score: 81

Sequence: 1 LLAGNSPRTQSPQNC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A Geneseq 16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	515	2	AAW54236 Human Lam
2	81	100.0	515	4	AB90770 Human ehe
3	81	100.0	664	6	ABU04992 Human exp
4	81	100.0	664	7	AD63975 Human pro
5	81	100.0	664	7	ADJ70273 Human hea
6	81	100.0	664	8	AD013765 Human wil
7	73	90.1	665	7	AD63973 Rat Prote
8	48	59.3	172	8	ADR10031 Human pro
9	46	56.8	103	4	AAW82436 Human imm
10	44	54.3	83	4	ABG22832 Novel hum
11	44	54.3	286	5	ABP43855 Mouse put
12	44	54.3	286	6	ADA55294 Human pro
13	44	54.3	443	4	AB63545 Drosophil
14	44	54.3	970	6	ABR53403 Protein s
15	44	54.3	970	7	ADK64724 Disease t
16	43	53.1	154	4	AAU52369 Propionib
17	43	53.1	154	6	ABM48888 Propionib
18	43	53.1	558	7	ADC08125 Rice prot
19	42	51.9	60	4	AAW88380 Human imm
20	42	51.9	105	2	AAW61569 Vpr prote
21	42	51.9	105	2	AAW68196 Vpr bindi
22	42	51.9	129	3	AAG10484 Arabidops
23	42	51.9	144	3	AAG10483 Arabidops
24	42	51.9	164	7	ABO81750 Pseudomon
25	42	51.9	177	2	AAW00642 Infectiou

26	42	51.9	177	2	AAW06794	ILTV univ
27	42	51.9	256	7	ADB65101	Human pro
28	42	51.9	613	3	AAG42554	Arabidops
29	42	51.9	684	3	AAG42553	Arabidops
30	42	51.9	690	4	ABB71234	Drosophil
31	42	51.9	735	4	ABG06831	Novel hum
32	42	51.9	735	7	ADC33376	Human nov
33	42	51.9	772	3	AAG42552	Arabidops
34	42	51.9	772	8	ADO61607	Transcrip
35	42	51.9	1224	6	ABR55649	Amino aci
36	42	51.9	2150	5	AAO22566	Wooden le
37	41	50.6	10	2	AAAR15781	Farnesyl-
38	41	50.6	10	2	AAAR49742	Farnesyl-
39	41	50.6	10	2	AAAR77804	Farnesyl
40	41	50.6	10	2	AAW04434	Farnesyl
41	41	50.6	68	4	AAU53546	Propionib
42	41	50.6	68	6	ABM50065	Propionib
43	41	50.6	539	7	ADL06330	Amorpha f
44	41	50.6	542	2	AAW26130	Amorpha f
45	41	50.6	747	8	ADR86215	Aspergill
46	41	50.6	791	8	ADQ66060	Novel hum
47	41	50.6	871	8	ADK52121	Human ato
48	41	50.6	871	8	ADR09093	Human pro
49	41	50.6	1937	4	ABR58985	Drosophil
50	40	49.4	71	4	ABG20384	Novel hum
51	40	49.4	82	4	AAU61331	Novel hum
52	40	49.4	82	6	ABM57850	Propionib
53	40	49.4	105	6	AAU66847	Propionib
54	40	49.4	105	6	ABM63366	Propionib
55	40	49.4	141	4	AAW86096	Human imm
56	40	49.4	170	7	ADM05956	Human pro
57	40	49.4	172	4	AAU55863	Propionib
58	40	49.4	172	6	ABM52382	Propionib
59	40	49.4	234	7	ADM04835	Human pro
60	40	49.4	452	4	ABR71874	Drosophil
61	40	49.4	462	4	AAU93757	Human pro
62	40	49.4	462	4	AAU95328	Human pro
63	40	49.4	462	4	AAU95328	Human pro
64	40	49.4	583	4	AAW40015	Human mem
65	40	49.4	599	3	AAW29657	Human mem
66	40	49.4	713	7	ADJ69809	Human hea
67	40	49.4	759	6	ABM68367	Phototrab
68	40	49.4	828	7	ADJ70343	Human hea
69	40	49.4	1020	6	ABR41949	Rat ATPas
70	40	49.4	1020	8	ADI24515	Human mod
71	40	49.4	1020	8	ADO28693	Human Na/
72	40	49.4	1054	8	ABO85034	Human can
73	40	49.4	1110	7	ADI21164	Novel hum
74	40	49.4	2063	8	ABO85033	Murine ca
75	39.5	48.8	130	2	AAU21425	Human hig
76	39.5	48.8	5877	6	AAE34702	Human muc
77	39.5	48.8	10431	6	ABU54861	Human CA1
78	39.5	48.8	22152	8	ABU72871	Human ova
79	39.5	48.8	22157	8	ADP84155	Human CA1
80	39	48.1	45	8	ABO56724	Human gen
81	39	48.1	57	4	AAU66292	Propionib
82	39	48.1	57	4	AAU46738	Propionib
83	39	48.1	57	6	ABM62811	Propionib
84	39	48.1	57	6	ABM43257	Propionib
85	39	48.1	81	5	ABP04444	Human ORF
86	39	48.1	105	5	ABP03018	Human ORF
87	39	48.1	110	5	ABP74902	Human NOV
88	39	48.1	114	5	ABP70134	Human NOV
89	39	48.1	144	7	ABO77040	Pseudomon
90	39	48.1	178	6	ABM62278	Propionib
91	39	48.1	267	3	AAU70223	Human RNA
92	39	48.1	267	3	AAU70223	Human RNA
93	39	48.1	267	4	AAU70223	Human RNA
94	39	48.1	281	2	AAU70223	Human RNA
95	39	48.1	341	7	ADC71244	Human col
96	39	48.1	362	8	ADM90321	Human pha
97	39	48.1	385	5	ABP41904	Human ova
98	39	48.1	399	2	AAU223846	Prepro fo

99 39 48.1 402 2 AAR27291 Human ost  
100 39 48.1 402 2 AAR44759 Human ost

## ALIGNMENTS

RESULT 1  
AAW54236  
ID AAW54236 standard; peptide; 515 AA.

AC AAW54236;  
XX  
XX 10-AUG-1998 (first entry)  
XX Human Lamin A protein.  
DE  
DE Cervical cancer-associated protein; Cvc; tryptic peptide; human;  
XX detection; treatment; Lamin A; non-chromatin protein.  
XX

OS Homo sapiens.

XX WO9809170-A2.

XX 05-MAR-1998.

XX PF 19-AUG-1997; 97WO-US014526.

XX PR 30-AUG-1996; 96US-00705660.

XX (MATR-) MATRITECH INC.

XX PI Keesee SK, Obar R, Wu Y;

XX DR WPI; 1998-230271/20.

XX Detection and therapy of cervical cancer - using specific cervical cancer  
PT -associated proteins as targets for treatment or as indicators for  
PT detection.

XX Claim 18; Page 68-69; 79pp; English.

XX This protein is the human lamin A which is used to obtain tryptic  
CC peptides which are used in a method for detecting cervical cancer. The  
CC method involves detecting the presence of a cervical cancer-associated  
CC protein (Cvc) in a tissue or body fluid sample. The Cvc is characterised  
CC as having a molecular weight of 44900-69400 Daltons as determined by  
CC sodium dodecyl-sulphate (SDS)-PAGE techniques and an isoelectric point  
CC (pI) of 5.1-6.6 as determined by standard isoelectric focusing  
CC techniques. The protein is further characterised as being a non-chromatin  
CC protein which is detectable at a higher level in a human cervical cancer  
CC cell than in a normal human cervical cell, as determined by 2D-gel  
CC electrophoresis. The methods can be used for the early and rapid  
CC detection of cervical cancer, for treating cervical cancers and for  
CC monitoring the efficacy of such treatment

XX Sequence 515 AA;

Query Match 100.0%; Score 81; DB 2; Length 515;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
Db 498 LLGNSSPRTQSPQNC 512

RESULT 2  
AAB90770  
ID AAB90770 standard; protein; 515 AA.

XX AAB90770;

XX

DT 15-JUN-2001 (first entry)  
XX Human shear stress-response protein SEQ ID NO: 40.  
DE Human; shear stress-response protein; vascular disease; arteriosclerosis.  
XX Homo sapiens.  
XX WO200125427-A1.  
XX 12-APR-2001.  
XX 02-OCT-2000; 2000WO-JP006840.  
XX 01-OCT-1999; 99JP-00280976.  
XX (KYOWA) KYOWA HAKKO KOGYO KK.  
XX (NOJI/) NOJIMA H.  
XX Nojima H, Yoshisue H, Obayashi M, Oka T, Kawabata A, Sakurada K;  
PI Kuga T, Sekine S, Nakamura Y, Sugano S;  
XX WPI; 2001-266308/27.  
XX N-PSDB; AAO02893.  
XX DNA sequences, proteins encoded by them and antibodies against them  
PT useful in diagnosis and treatment of vascular disease caused by  
PT arteriosclerosis.  
XX Claim 60; Page 293-296; 678pp; Japanese.  
XX The present invention provides the protein and coding sequences of a  
CC number of human shear stress response proteins. These are useful in the  
CC diagnosis, treatment and screening of vascular diseases caused by  
CC arteriosclerosis, including heart failure, post-PTCA restenosis and  
CC hypertension  
XX Sequence 515 AA;

Query Match 100.0%; Score 81; DB 4; Length 515;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
Db 498 LLGNSSPRTQSPQNC 512

RESULT 3  
ABU04992  
ID ABU04992 standard; protein; 664 AA.

XX ABU04992;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1658.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX protease; protease inhibitor; transporter; cytoskeletal protein;  
XX receptor; transcription factor; cancer; MHC;  
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

```
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0359985P.
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1658; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (BPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 664 AA;
XX
XX Query Match 100.0%; Score 81; DB 6; Length 664;
XX Best Local Similarity 100.0%; Pred. No. 0.00015;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LLGNSSPRTQSPQNC 15
XX DB 647 LLGNSSPRTQSPQNC 661
XX
XX RESULT 4
XX ADE63975
XX ID ADE63975 standard; protein; 664 AA.
XX AC ADE63975;
XX AC
XX DT 29-JAN-2004 (first entry)
XX XX
XX DE Human Protein P02545, SEQ ID NO 9921.
XX
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX
XX XX Homo sapiens.
XX OS
XX PN WO2003016475-A2.
XX XX
XX PD 27-FEB-2003.
XX XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX
```

```
XX (GEHO ) GEN HOSPITAL CORP.
XX (FARB ) BAYER AG.
XX
XX WPI; 2003-268312/26.
XX GENBANK; P02545.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 664 AA;
XX
XX Query Match 100.0%; Score 81; DB 7; Length 664;
XX Best Local Similarity 100.0%; Pred. No. 0.00015;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LLGNSSPRTQSPQNC 15
XX DB 647 LLGNSSPRTQSPQNC 661
XX
XX RESULT 5
XX ADJ70273
XX ID ADJ70273 standard; protein; 664 AA.
XX XX
XX AC ADJ70273;
XX XX
XX DT 06-MAY-2004 (first entry)
XX XX
XX DE Human heat mitochondrial protein as a therapeutic target SeqID2079.
XX
XX KW mitochondrial; human; screening assay; diabetes mellitus;
XX KW Huntington's disease; osteoarthritis;
XX KW Leber's hereditary optic neuropathy; LHON;
XX KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
XX KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
XX KW osteopathic; ophthalmological; cytostatic.
XX
XX OS Homo sapiens.
XX
```

```

PN WO2003087768-A2.
XX
PD
XX
XX
XX PF 23-OCT-2003.
XX
XX
XX PF 04-APR-2003; 2003WO-US010870.
XX
XX PR 12-APR-2002; 2002US-0372843P.
XX
XX PR 17-JUN-2002; 2002US-0399987P.
XX
XX PR 20-SEP-2002; 2002US-0412418P.
XX
XX
XX PA (MITO-) MITOKOR.
XX
XX PA (BUCK-) BUCK INST AGE RES.
XX
XX GHOSH SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
PI Warnock DS;
XX
XX WPI; 2003-845369/78.
XX
XX PT Identifying a mitochondrial target for drug screening assays and for
XX treating diseases associated with altered mitochondrial function,
XX PT comprises detecting a modified polypeptide in a sample and correlating
XX PT with the disease.
XX
XX PS Claim 1; SEQ ID NO 2079; 180pp; English.
XX
XX CC This invention relates to novel mitochondrial targets that can be used
XX for therapeutic intervention in treating a disease associated with
XX CC altered mitochondrial function. Specifically, it refers to a method for
XX CC identifying proteins of the human heart mitochondrial proteome that are
XX CC useful for drug screening assays, as well as therapeutic targets. The
XX CC present invention describes a method for identifying such proteins that
XX CC can be used in the treatment of various diseases associated with altered
XX CC mitochondrial function including diabetes mellitus, Huntington's disease,
XX CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX CC encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX CC compositions have neuroprotective, neurotropic, antidiabetic,
XX CC anticonvulsant, antiarrhythmic, osteopathic, ophthalmological and
XX CC cytoskeletal activities. This polypeptide sequence is a human heart
XX CC mitochondrial protein of the invention.
XX
XX SQ Sequence 664 AA;

Query Match 100.0%; Score 81; DB 7; Length 664;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQNC 15
Db |||||
647 LLGNSSPRTQSPQNC 661

RESULT 6
AD013765
ID AD013765 standard; protein; 664 AA.
XX
XX AC AD013765;
XX
XX DT 15-JUL-2004 (first entry)
XX
XX DE Human wild type Laminin A (LMNA) protein.
XX
XX KW antiarteriosclerotic; laminin A; mutation; diagnosis; progeroid disease;
XX KW Hutchinson-Gilford Progeria Syndrome; arteriosclerosis; atherosclerosis.
XX
XX OS Homo sapiens.
XX
XX FN WO2004035753-A2.
XX
XX PD 29-APR-2004.
XX
XX PF 17-OCT-2003; 2003WO-US033058.
XX

18-OCT-2002; 2002US-0419541P.
14-APR-2003; 2003US-0463084P.
XX
XX PA (PROG-) PROGERIA RES FOUND INC.
XX PA (NYME-) NEW YORK STATE OFFICE MENTAL HEALTH.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Eriksson MBH, Collins FS, Gordon LB, Brown TW;
XX DR WPI; 2004-348447/32.
XX DR N-PSDB; ADO13764.
XX
XX PT Detecting a biological condition associated with a dominant laminin A
XX (LMNA) mutation, useful for diagnosing, preventing or treating a
XX PT progeroid disease that is Hutchinson-Gilford Progeria Syndrome, and/or
XX PT arteriosclerosis.
XX
XX PS Disclosure; SEQ ID NO 2; 85pp; English.
XX
XX CC The invention relates to a method of detecting a biological condition
XX CC associated with a dominant laminin A (LMNA) mutation in a subject
XX CC comprising determining whether a subject has mutation in LMNA, and where
XX CC the mutation comprises a variant nucleic acid sequence in or
XX CC corresponding to codon 608, 644, 145, 471, 527 or 269 of human LMNA, or
XX CC two or more mutations. The methods and compositions of the present
XX CC invention are useful for the diagnosis, prevention and/or treatment of
XX CC diseases or conditions associated with the mutation of LMNA, such as
XX CC progeroid disease that is Hutchinson-Gilford Progeria Syndrome, or
XX CC arteriosclerosis or atherosclerosis. This sequence corresponds to the
XX CC wild type laminin A protein.
XX
XX SQ Sequence 664 AA;

Query Match 100.0%; Score 81; DB 8; Length 664;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQNC 15
Db |||||
647 LLGNSSPRTQSPQNC 661

RESULT 7
ADE63973
ID ADE63973 standard; protein; 665 AA.
XX
XX AC ADE63973;
XX
XX DT 29-JAN-2004 (first entry)
XX
XX DE Rat Protein P48679, SEQ ID NO 9919.
XX
XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX OS Rattus norvegicus.
XX
XX FN WO2003016475-A2.
XX
XX PD 27-FEB-2003.
XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX
XX PR 01-NOV-2001; 2001US-0346382P.
XX
XX PR 26-NOV-2001; 2001US-0333347P.
XX
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX DR WPI; 2003-268312/26.

```

DR GENBANK; P48679.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX

XX Claim 1; Page: 1017pp; English.

XX

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a

CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the

CC activity in an animal of one or more of the polypeptides given in the

CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more

CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

CC therapy). The sequence presented is a rat protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

XX SQ Sequence 665 AA;

Query Match 90.1%; Score 73; DB 7; Length 665;

Best Local Similarity 93.3%; Pred. NO. 0.0032;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LLGNSSPRTQSPQC 15

DB 648 LLGNSSPRTQSQNC 662

RESULT 8

ID ADR10031 standard; protein; 172 AA.

XX

XX ADR10031;

XX

XX 04-NOV-2004 (first entry)

XX

XX Human protein useful for treating neurological disease Seq 3537.

XX

XX human; oligo-capping method; diagnostic marker; gene therapy;

XX osteoporosis; neurological disease; Alzheimer's disease;

XX Parkinson's disease; dementia; short memory; cancer;

XX sense or motor function; emotional reaction; fear response; panic;

XX osteopathic; neuroprotective; nootropic; antiparkinsonian; cytotstatic;

XX tranquiliser.

XX

XX Homo sapiens.

XX

XX EP1447413-A2.

XX

XX 18-AUG-2004.

XX

XX 12-FEB-2004; 2004EP-00003145.

XX

XX 14-FEB-2003; 2003JP-00102207.

XX

XX 09-MAY-2003; 2003JP-00131452.

XX

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;

PI Wakamatsu A, Ishii S, Nagai K, Irie R;

XX

XX WPI; 2004-583265/57.

DR N-PSDB; ADR08075.

XX

XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,

PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX

XX Claim 1; SEQ ID NO 3537; 2686pp; English.

XX

XX This invention relates to novel, isolated full length human cDNA

CC molecules and the encoded proteins thereof. Specifically, it refers to

CC cDNA clones obtained by an oligo-capping method, where none of these

CC clones are identical to any known human mRNAs. The present invention

CC describes an immunoassay to identify agonists and antagonists, as well as

CC antibodies, antisense molecules and siRNAs that can all be used to bind

CC to and modulate expression of the cDNA molecules. As such, these

CC molecules are useful for diagnostic markers or therapeutic targets for

CC the various diseases or morbid states. In particular, they are useful in

CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's

CC disease, Parkinson's disease, dementia, short memory and various cancers,

CC as well as for maintaining equilibrium of sense or motor function, and

CC for treating emotional reaction, fear response and panic. Accordingly,

CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,

CC cytosstatic and tranquiliser activities. This polypeptide is a protein

CC encoded by a full length human cDNA sequence of the invention. NOTE: This

CC sequence is not given in the sequence listing of the specification but

CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-

CC office.

XX

XX SQ Sequence 172 AA;

Query Match 59.3%; Score 48; DB 8; Length 172;

Best Local Similarity 76.9%; Pred. No. 10;

Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LLGNSSPRTQSPQ 13

DB 108 VLGYSPPRTLSPO 120

RESULT 9

AA82436

ID AA82436 standard; protein; 103 AA.

XX

XX AA82436;

XX

XX 07-NOV-2001 (first entry)

XX

XX Human immune/haematopoietic antigen SEQ ID NO:10029.

XX

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytotstatic; gene therapy; vaccine; metastasis.

XX

XX Homo sapiens.

OS

WO200157182-A2.

PN

XX

XX 09-AUG-2001.

XX

XX 17-JAN-2001; 2001WO-US001354.

XX

XX 31-JAN-2000; 2000US-0179065P.

PR

XX 04-FEB-2000; 2000US-0180628P.

PR

XX 24-FEB-2000; 2000US-0184664P.

PR

XX 02-MAR-2000; 2000US-0186350P.

PR

XX 16-MAR-2000; 2000US-0189874P.

PR

XX 17-MAR-2000; 2000US-0190076P.

PR

XX 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0227009P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 27-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 08-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
XX N-PSDB; AAK55217.  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX Claim 11; SEQ ID NO 10029; 3071pp + Sequence Listing; English.  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
XX amino acid sequences given in AAK6170 to AAK61921. (I) have cytostatic  
XX activity, and can be used in gene therapy and vaccine production. (I)  
XX proteins and polynucleotides may be used in the prevention, diagnosis and  
XX

CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention

XX  
 XX  
 SQ Sequence 103 AA;

Query Match 56.8%; Score 46; DB 4; Length 103;  
 Best Local Similarity 57.1%; Pred. NO. 13;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LGNSPRTQSPQNC 15  
 ||: :| ||| |  
 Db 71 LGSWTPNTQPPQRC 84

RESULT 10  
 ID ABG22832 standard; protein; 83 AA.  
 XX  
 AC ABG22832;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #22823.  
 XX  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 KW  
 KW Homo sapiens.  
 XX  
 OS WO200175067-A2.  
 XX  
 PN 11-OCT-2001.  
 XX  
 PD 30-MAR-2001; 2001WO-US008631.  
 XX  
 PF 31-MAR-2000; 2000US-00540217.  
 XX  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Drmanac RT, Liu C, Tang YT;  
 PI  
 XX WPI; 2001-639362/73.  
 XX  
 DR N-PSDB; AAS87019.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PT  
 XX Claim 20; SEQ ID NO 53191; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX SQ Sequence 83 AA;

Query Match 54.3%; Score 44; DB 4; Length 83;  
 Best Local Similarity 80.0%; Pred. No. 21;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NSSPRTQSPQ 13  
 |||||:| |:  
 Db 4 NSSPRSQSPR 13

RESULT 11  
 ID ABP43855 standard; protein; 286 AA.  
 XX  
 AC ABP43855;  
 XX  
 DT 26-FEB-2003 (first entry)  
 XX  
 DE Mouse putative protein #15.  
 XX  
 DE Neuroprotective; immunomodulator; cancer; cytostatic; anti-inflammatory;  
 KW gene therapy; nutritional supplement; wound; burn; ulcer;  
 KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;  
 KW autoimmune disorder; inflammation; vulnery.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200231111-A2.  
 XX  
 PD 18-APR-2002.  
 XX  
 PF 11-OCT-2001; 2001WO-US027760.  
 XX  
 PR 12-OCT-2000; 2000US-00687527.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
 XX  
 DR WPI; 2002-426278/45.  
 XX  
 DR N-PSDB; ABQ61099.  
 XX  
 XX New polypeptides and their encoded proteins, useful as nutritional  
 PT sources or supplements, or in gene therapy, particularly for treating  
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or  
 PT inflammation.  
 PT  
 XX Claim 20; SEQ ID # 758; 357pp + Sequence Listing; English.  
 PS  
 XX The invention relates to 446 newly isolated polynucleotide sequences. The  
 CC activity of polynucleotides of the invention may be described as,  
 CC vulnery, neuroprotective, immunomodulator, cytostatic and anti-  
 CC inflammatory. Compositions comprising nucleic acids of the invention are  
 CC useful for treating a mammalian subject, or as nutritional sources or  
 CC supplements. These are useful in gene therapy, particularly for treating  
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,  
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or  
 CC inflammation. The nucleic acids and polypeptides are also useful in  
 CC diagnostic and research methods. The sequences given in records ABP43544-





```

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59554.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 13564; 1069pp; English.
XX
XX Sequences AAU39105-AAUG8017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 154 AA;
XX
XX Query Match 53.1%; Score 43; DB 4; Length 154;
XX Best Local Similarity 58.3%; Pred. No. 61;
XX Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 NSSPRTQSPQNC 15
Db :|||:|:|:|
13 SSSPQTSSPKRC 24

RESULT 17
ABM48888
ID ABM48888 standard; protein; 154 AA.
AC ABM48888;
XX
XX 20-OCT-2003 (first entry)
XX
XX Propionibacterium acnes predicted ORF-encoded polypeptide #13564.
DE
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
XX Propionibacterium acnes.
OS
XX
XX WO2003033515-A1.
PN
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
PF
XX
XX 15-OCT-2001; 2001US-00978825.
PR
XX
XX (CORI-) CORIYA CORP.
PA
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
XX WPI; 2003-381789/36.
DR

```

---

```

DR N-PSDB; ACF64483.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 13564; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 154 AA;
XX
XX Query Match 53.1%; Score 43; DB 6; Length 154;
XX Best Local Similarity 58.3%; Pred. No. 61;
XX Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 NSSPRTQSPQNC 15
Db :|||:|:|:|
13 SSSPQTSSPKRC 24

RESULT 18
ADC08125
ID ADC08125 standard; protein; 558 AA.
XX
XX ADC08125;
XX
XX 18-DEC-2003 (first entry)
XX
XX Rice protein sequence Seq ID392 related to grain filling.
DE
XX
XX plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
XX carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
XX tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
XX wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
XX gene; ds; plant.
XX
XX Oryza sativa.
OS
XX
XX WO2003000905-A2.
PN
XX
XX 03-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-IB002450.
XX
XX 22-JUN-2001; 2001US-0300112P.
XX
XX 26-SEP-2001; 2001US-0325277P.
XX

```



PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0251989P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
XX N-PSDB; AAK61161.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX Claim 11; SEQ ID NO 15973; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
XX activity, and can be used in gene therapy and vaccine production. (I)  
XX proteins and polynucleotides may be used in the prevention, diagnosis and  
XX treatment of diseases associated with inappropriate (I) expression. For  
XX example, they may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
XX SQ Sequence 60 AA;  
  
XX Query Match 51.9%; Score 42; DB 4; Length 60;  
XX Best Local Similarity 64.3%; Pred. No. 32;  
XX Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
XX QY 1 LLGNSPRTQSPQN 14  
XX |||:|||||  
XX Db 34 LLGSSPRTSVSQN 47  
  
XX RESULT 20  
XX AAW61569  
XX ID AAW61569 standard; peptide; 105 AA.  
XX AC AAW61569;  
XX XX 27-OCT-1998 (first entry)  
XX DX Vpr protein binding B29-1 amino acid sequence.  
XX DE Vpr protein binding B29-1 amino acid sequence.  
XX XX Lentiviral infection; Vpr protein; HIV infection; cell stasis;  
XX KW cell death.  
XX OS Homo sapiens.  
XX XX WO9835234-A1.  
XX XX 13-AUG-1998.  
XX XX 11-FEB-1998; 98WO-US003008.  
XX PF 11-FEB-1997; 97US-00797907.  
XX PR (REGC ) UNIV CALIFORNIA.  
XX XX Chen ISY, Jowett JBM, Withers-Ward E;  
XX XX WPI; 1998-447375/38.  
XX XX Identification of compounds binding the HIV-1 Vpr protein - that block  
XX PT Vpr-mediated cell stasis, useful for treating HIV-1 or other lentiviral  
XX PT infections.  
XX XX Disclosure; Fig 8; 63pp; English.  
XX XX  
XX This represents the amino acid sequence of the B29-1 protein. This  
XX protein can bind to the Vpr protein encoded by the HIV genome. The  
XX invention provides a method of identifying an agent for use in treating  
XX lentiviral infections. The method comprises contacting a cellular target  
XX of the Vpr protein with the agent to be tested, and assessing the ability  
XX of the agent to block interaction of the Vpr protein with the cellular  
XX target, where an agent which blocks this interaction is an anti-  
XX lentiviral agent. Alternatively, the agent contacts a cell expressing the  
XX Vpr protein under conditions where the Vpr protein induces cell stasis in  
XX the absence of the agent. Identification of the agent is then observed by  
XX blockage of Vpr-induced cell stasis. The method allows the identification  
XX of compounds that block Vpr-mediated cell stasis and ultimately cell  
XX death. The compounds can thus be used in the treatment of HIV and other  
XX lentiviral infections

```
SQ Sequence 105 AA;
Query Match 51.9%; Score 42; DB 2; Length 105;
Best Local Similarity 80.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GNSSPRTQSP 12
DB 83 GNSSPHTASP 92

RESULT 21
AAW68196
ID AAW68196 standard; peptide; 105 AA.
XX
AC AAW68196;
XX
DT 25-MAR-2003 (revised)
DT 29-OCT-1998 (first entry)
XX
DE Vpr binding protein B29-1 amino acid sequence.
XX
KW Lentivirus; Vpr protein; HIV infection; cell stasis; cell death; cancer;
KW autoimmune disease; B29-1.
XX
OS Homo sapiens.
XX
PN WO9835032-A2.
XX
PD 13-AUG-1998.
XX
PF 11-FEB-1998; 98WO-US003390.
XX
PR 11-FEB-1997; 97US-00798597.
XX
PR 24-OCT-1997; 97US-00959279.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Chen ISY, Jowett JEM, Withers-Ward E, Stewart SA, Poon B;
PI Feigson J, Dieckmann T;
XX
XX WPI; 1998-447229/38.
XX
PT Arresting cell growth using lentivirus Vpr virion protein - used for
PT treatment of cancer and screening for agents that reduce Vpr binding,
PT e.g. anti-HIV agents.
XX
PS Disclosure; Fig 7; 71pp; English.
XX
CC This represents the amino acid sequence of the B29-1 protein. This
CC protein can bind to the Vpr protein encoded by the HIV genome. This is
CC used as a cellular target in the method of the invention of identifying
CC antitumour therapeutic candidates. The invention provides a method for
CC arresting the growth of a cell by treatment with a Vpr lentivirus protein
CC or its analogue. Agents that reduce binding of Vpr to a cellular target
CC are useful for treating HIV (human immune deficiency virus) infection or
CC more generally for restoring growth. The antitumour agent identified is
CC useful for treating any type of cancer, since it induces cell stasis
CC (blocks development at the G2 stage) and death. The agents can also be
CC used for treating autoimmune diseases. (Updated on 25-MAR-2003 to correct
CC PI field.)
XX
SQ Sequence 105 AA;
Query Match 51.9%; Score 42; DB 2; Length 105;
Best Local Similarity 80.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GNSSPRTQSP 12
DB 83 GNSSPHTASP 92
```

---

```
RESULT 22
AAG10484
ID AAG10484 standard; protein; 129 AA.
XX
AC AAG10484;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8826.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127482P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
```

PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 28-SEP-1999; 99US-0155659P.  
PR 29-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160801P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.9%; Score 42; DB 3; Length 129;  
Best Local Similarity 61.5%; Pred. No. 74;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LGNSSPRTQSPQN 14  
|||||:|  
Db 24 LGNSAPISPPN 36

RESULT 23  
RAG10483  
ID AAG10483 standard; protein; 144 AA.  
XX AAG10483; AC  
XX AAG10483; AC  
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 8825.  
DE  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0137828P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140655P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144844P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 23-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.

```

PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.9%; Score 42; DB 3; Length 144;
Best Local Similarity 61.5%; Pred. No. 83;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LGNSPRTQSPQN 14
DB 39 LGNSAPISSPPN 51

RESULT 24
ABO81750
ID ABO81750 standard; protein; 164 AA.
AC ABO81750;
XX
XX 29-JUL-2004 (first entry)
DT
DE Pseudomonas aeruginosa polypeptide #13925.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX

```

```

PN US6551795-B1.
XX
PD 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX N-PSDB; ABD15321.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 30496; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 164 AA;

Query Match 51.9%; Score 42; DB 7; Length 164;
Best Local Similarity 63.6%; Pred. No. 96;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 SSPRTQSPQNC 15
DB 22 SNPRHQEPQSC 32

RESULT 25
AAW00642
ID AAW00642 standard; protein; 177 AA.
XX
XX AAW00642;
XX
XX 16-OCT-2003 (revised)
DT 19-NOV-1996 (first entry)
XX
XX Infectious laryngotracheitis virus ORF8 RC product.
XX
XX Infectious laryngotracheitis virus; ILTV; herpesvirus; attenuation;
XX vector; vaccine; chicken; poultry; immunisation.
XX
XX Gallid herpesvirus 1.
XX OS
XX WO9508622-A1.
XX
XX 30-MAR-1995.
PD
XX
XX 16-SEP-1994; 94WO-US010628.
XX PF
XX
XX 24-SEP-1993; 93US-00126597.
XX PR

```

```

XX (SYTR ) SYNTRO CORP.
XX PA
XX PI Wild MA, Cochran MD;
XX DR WPI; 1995-139591/18.
XX DR N-PSDB; AAT33513.
XX PT Recombinant attenuated infectious laryngotracheitis virus - for use in
XX PT vaccines to protect poultry from infection from the virus, also methods
XX PT of distinguishing between vaccinated and naturally infected birds.
XX PS Example 1; Page 116-117; 177pp; English.
XX CC A protein (AAW06794) of unknown function is the product of the ORF8
XX CC reverse complement gene (AAT33513) transcribed from bases 11,150-10,617
XX CC on the reverse complement strand of the unique short region (see also
XX CC AAT33504) of infectious laryngotracheitis virus (ILTV) genomic DNA. The
XX CC protein has a mol.wt. of approx. 19,470. (Updated on 16-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 177 AA;

  Query Match      51.9%; Score 42; DB 2; Length 177;
  Best Local Similarity 53.3%; Pred. No. 1e+02;
  Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQNC 15
   |||||
DB 5 LLNRGSPRLNSPPKC 19

RESULT 26
AATW06794
ID AAW06794 standard; protein; 177 AA.
XX AC AAW06794;
XX DT 16-OCT-2003 (revised)
XX DT 02-JUN-1997 (first entry)
XX DE ILTV unique short region ORF8 (reverse complement) product.
XX KW ILTV; vaccine; vector; attenuation; poultry;
XX KW avian infectious bronchitis virus; Newcastle disease virus;
XX KW infectious bursal disease virus of chickens; Marek's disease virus;
XX KW herpesvirus.
XX OS Gallid herpesvirus 1; USDA strain 8302.
XX PN WO9629396-A1.
XX PD 26-SEP-1996.
XX PF 21-MAR-1996; 96WO-US003916.
XX PR 23-MAR-1995; 95US-00410121.
XX PR 06-JUN-1995; 95US-00468190.
XX PA (SYTR ) SYNTRO CORP.
XX PI Wild MA, Cochran MD;
XX DR WPI; 1996-443172/44.
XX DR N-PSDB; AAT44384, AAT44385.
XX PT Recombinant infectious laryngotracheitis virus with deletion in the
XX PT glyco:protein G, GI or US2 gene, etc. - useful for vaccines against
XX PT infectious laryngotracheitis in poultry.
XX PS Example 11; Page 125; 216pp; English.
XX CC A 177-amino acid polypeptide (AAW06794) of unknown function is encoded by

```

---

```

CC open reading frame 8 (reverse complement) of the unique short region
CC (AAT44384) of infectious laryngotracheitis virus (ILTV). Attenuated
CC recombinant ILTVs have at least one deletion in the unique short region
CC and can be used in vaccines against ILTV, and also in multivalent
CC vaccines for one or more avian diseases. (Updated on 16-OCT-2003 to
CC standardise OS field)
XX SQ Sequence 177 AA;

  Query Match      51.9%; Score 42; DB 2; Length 177;
  Best Local Similarity 53.3%; Pred. No. 1e+02;
  Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQNC 15
   |||||
DB 5 LLNRGSPRLNSPPKC 19

RESULT 27
ADB65101
ID ADB65101 standard; protein; 256 AA.
XX AC ADB65101;
XX DT 04-DEC-2003 (first entry)
XX DE Human protein encoded by clone SPLEN20015100.
XX KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
XX KW cell regeneration; membrane protein; signal transduction-related protein;
XX KW transcription-related protein; osteoporosis; neurological disease;
XX KW cancer; tumour.
XX OS Homo sapiens.
XX PN EP1308459-A2.
XX PD 07-MAY-2003.
XX PF 28-MAR-2002; 2002EP-00007401.
XX PR 05-NOV-2001; 2001JP-00379298.
XX PR 25-JAN-2002; 2002US-00350978.
XX PA (HELI-) HELIX RES INST.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX DR WPI; 2003-450961/43.
XX DR N-PSDB; ADB63131.
XX PT New polynucleotides and polypeptides, useful for developing a diagnostic
XX PT marker or medicines for regulation of their expression and activity, or
XX PT as targets of gene therapy.
XX PS Claim 1; Page; 222pp; English.
XX CC The invention discloses a polynucleotide comprising a sequence selected
XX CC from 1970 fully defined nucleotide sequences which encode novel
XX CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
XX CC or its partial peptide, an antibody binding to the polypeptide or peptide
XX CC of the polynucleotide, immunologically assaying the polypeptide or
XX CC peptide of the polynucleotide by contacting the polypeptide or peptide
XX CC with the antibody of the encoded protein, and observing the binding
XX CC between the two, a transformant carrying the polynucleotide in an
XX CC expressible manner and an antisense polynucleotide. The oligonucleotide
XX CC is useful as a primer for synthesising the polynucleotide, or as a probe
XX CC for detecting the polynucleotide. The polynucleotides and encoded
XX CC proteins are useful as pharmaceutical agents and many disease-related
XX CC genes may be included in them, for developing a diagnostic marker or

```

CC medicines for regulation of their expression and activity, or as targets  
CC of gene therapy. The genes are involved in tissue and/or cell  
CC regeneration. Membrane proteins, signal transduction-related proteins,  
CC transcription-related proteins, disease-related proteins and genes  
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,  
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate  
CC the activity or expression of the encoded protein to treat diseases. The  
CC sequence presented is a protein of the invention. Note: Some of the  
CC sequence data for this patent is not represented in the printed  
CC specification, but is based on sequence information supplied by the  
CC European Patent Office.

XX Sequence 256 AA;

SQ Query Match 51.9%; Score 42; DB 7; Length 256;

Best Local Similarity 80.0%; Pred. No. 1.5e-02;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GNSSPRTQSP 12

|||||

Db 234 GNSSPHTASP 243

RESULT 28

AAG42554

ID AAG42554 standard; protein; 613 AA.

XX AC AAG42554;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 53084.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.

```
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149360P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149428P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.

PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.9%; Score 42; DB 3; Length 613;
Best Local Similarity 61.5%; Pred. No. 3.9e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LGNSPRTQSPQN 14
    |||||:|:|
Db 508 LGNSAPISSPPN 520

RESULT 29
AAG42553
ID AAG42553 standard; protein; 684 AA.
AC AAG42553;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53083.
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135623P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
```

PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137503P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 21-JUN-1999; 99US-013989P.  
PR 22-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 23-JUN-1999; 99US-0140355P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-014230P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-014292P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144633P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151106P.  
PR 27-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 05-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

```

Query Match          51.9%; Score 42; DB 3; Length 684;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LGNSSPRTQSPQN 14
DB 579 LGNSAPISSPPN 591

RESULT 30
ABB71234
ID ABB71234 standard; protein; 690 AA.
XX
AC ABB71234;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 40494.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
DR N-PSDB; ABLI5337.
XX
New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions.
XX
PS Disclosure; SEQ ID NO 40494; 21pp + Sequence Listing; English.
XX
The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
sequences (ABLI01840-ABLI6175) and the encoded proteins (ABB57737-
ABB72072). The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 690 AA;

Query Match          51.9%; Score 42; DB 4; Length 690;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LGNSSPRTQSPQN 14
DB 55 VIGNSLTSPAPQN 68

RESULT 31
ABG06831
ID ABG06831 standard; protein; 735 AA.
XX
AC ABG06831;

```

```

XX 13-FEB-2002 (first entry)
DT Novel human diagnostic protein #6822.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS71018.
XX New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.
XX Claim 20; SEQ ID NO 37190; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical imaging
of sites expressing (II). (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
amino acid sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 735 AA;

Query Match          51.9%; Score 42; DB 4; Length 735;
Best Local Similarity 57.1%; Pred. No. 4.8e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQN 14
DB 129 LLGSPVPAQSPQS 142

RESULT 32
ADC33376
ID ADC33376 standard; protein; 735 AA.
XX
AC ADC33376;
XX
DT 18-DEC-2003 (first entry)

```

XX DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3458.  
XX KW Human; diagnostic; drug screening; forensics; gene mapping;  
XX KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
XX KW ulcers; osteoporosis; autoimmune disease; cancer;  
XX KW molecular weight marker; food supplement; antiparkinsonian; nootropic;  
XX KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;  
XX KW antilulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
XX KW gene therapy; chromosome 17.  
XX OS Homo sapiens.  
XX FN WO2003029271-A2.  
XX PD 10-APR-2003.  
XX PF 24-SEP-2002; 2002WO-US030474.  
XX PR 24-SEP-2001; 2001US-0324631P.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
PI Haley-Vicente D, Drmanac RT;  
XX DR WPI; 2003-371981/35.  
XX DR N-PSDB; ADC32609.  
XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or  
XX PT treating conditions such as neurodegenerative diseases, anemias, platelet  
XX PT disorders, wounds, ulcers, burns, osteoporosis, autoimmune diseases or  
XX PT cancer.  
XX PS Example 2; SEQ ID NO 3458; 1185pp; English.  
XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31961-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a human contig-  
CC encoded polypeptide sequence used in an example of the invention. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX CC Sequence 735 AA;

Query Match 51.9%; Score 42; DB 7; Length 735;  
Best Local Similarity 57.1%; Pred. NO. 4.8e+02;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LLGNSSPRTQSPQN 14  
Db 129 LLGSPVPQAQSPQS 142  
|||||:|||||:  
RESULT 33  
AAG42552  
ID AAG42552 standard; protein; 772 AA.  
XX AC AAG42552;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 53082.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EF1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
XX PR 05-MAR-1999; 99US-0123180P.  
XX PR 09-MAR-1999; 99US-0123548P.  
XX PR 23-MAR-1999; 99US-0125788P.  
XX PR 25-MAR-1999; 99US-0126264P.  
XX PR 29-MAR-1999; 99US-0126785P.  
XX PR 01-APR-1999; 99US-0127462P.  
XX PR 06-APR-1999; 99US-0128234P.  
XX PR 08-APR-1999; 99US-0128714P.  
XX PR 16-APR-1999; 99US-0129845P.  
XX PR 19-APR-1999; 99US-0130077P.  
XX PR 21-APR-1999; 99US-0130449P.  
XX PR 23-APR-1999; 99US-0130510P.  
XX PR 28-APR-1999; 99US-0130891P.  
XX PR 30-APR-1999; 99US-0131499P.  
XX PR 30-APR-1999; 99US-0132407P.  
XX PR 04-MAY-1999; 99US-0132484P.  
XX PR 05-MAY-1999; 99US-0132485P.  
XX PR 06-MAY-1999; 99US-0132486P.  
XX PR 06-MAY-1999; 99US-0132487P.  
XX PR 07-MAY-1999; 99US-0132863P.  
XX PR 11-MAY-1999; 99US-0134256P.  
XX PR 14-MAY-1999; 99US-0134218P.  
XX PR 14-MAY-1999; 99US-0134219P.  
XX PR 14-MAY-1999; 99US-0134221P.  
XX PR 14-MAY-1999; 99US-0134370P.  
XX PR 18-MAY-1999; 99US-0134768P.  
XX PR 19-MAY-1999; 99US-0134941P.  
XX PR 20-MAY-1999; 99US-0135124P.  
XX PR 21-MAY-1999; 99US-0135353P.  
XX PR 25-MAY-1999; 99US-0135629P.  
XX PR 25-MAY-1999; 99US-0136021P.  
XX PR 27-MAY-1999; 99US-0136392P.  
XX PR 28-MAY-1999; 99US-0136782P.  
XX PR 01-JUN-1999; 99US-0137222P.  
XX PR 03-JUN-1999; 99US-0137528P.  
XX PR 04-JUN-1999; 99US-0137502P.  
XX PR 08-JUN-1999; 99US-0138094P.  
XX PR 10-JUN-1999; 99US-0138540P.  
XX PR 10-JUN-1999; 99US-0138847P.

PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139454P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139455P.	PR	16-AUG-1999;	99US-0149358P.
PR	18-JUN-1999;	99US-0139456P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139457P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139460P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139462P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139463P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139464P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139465P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139466P.	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	99US-0139467P.	PR	27-AUG-1999;	99US-0151080P.
PR	18-JUN-1999;	99US-0139468P.	PR	30-AUG-1999;	99US-0151303P.
PR	18-JUN-1999;	99US-0139469P.	PR	31-AUG-1999;	99US-0151388P.
PR	21-JUN-1999;	99US-0139817P.	PR	01-SEP-1999;	99US-0151930P.
PR	22-JUN-1999;	99US-0139899P.	PR	07-SEP-1999;	99US-0152363P.
PR	23-JUN-1999;	99US-0140353P.	PR	10-SEP-1999;	99US-0153070P.
PR	23-JUN-1999;	99US-0140354P.	PR	13-SEP-1999;	99US-0153758P.
PR	24-JUN-1999;	99US-0140698P.	PR	15-SEP-1999;	99US-0154018P.
PR	28-JUN-1999;	99US-0140823P.	PR	16-SEP-1999;	99US-0154039P.
PR	29-JUN-1999;	99US-0140991P.	PR	20-SEP-1999;	99US-0154779P.
PR	30-JUN-1999;	99US-0141287P.	PR	22-SEP-1999;	99US-0155139P.
PR	01-JUL-1999;	99US-0141842P.	PR	23-SEP-1999;	99US-0155486P.
PR	01-JUL-1999;	99US-0142154P.	PR	24-SEP-1999;	99US-0155659P.
PR	02-JUL-1999;	99US-0142058P.	PR	28-SEP-1999;	99US-0156458P.
PR	06-JUL-1999;	99US-0142390P.	PR	29-SEP-1999;	99US-0156596P.
PR	08-JUL-1999;	99US-0142803P.	PR	04-OCT-1999;	99US-0157117P.
PR	09-JUL-1999;	99US-0142920P.	PR	05-OCT-1999;	99US-0157753P.
PR	12-JUL-1999;	99US-0142977P.	PR	06-OCT-1999;	99US-0157865P.
PR	13-JUL-1999;	99US-0143542P.	PR	07-OCT-1999;	99US-0158029P.
PR	14-JUL-1999;	99US-0143624P.	PR	08-OCT-1999;	99US-0158232P.
PR	15-JUL-1999;	99US-0144005P.	PR	12-OCT-1999;	99US-0158369P.
PR	16-JUL-1999;	99US-0144085P.	PR	13-OCT-1999;	99US-0159293P.
PR	16-JUL-1999;	99US-0144086P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144325P.	PR	14-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144331P.	PR	14-OCT-1999;	99US-0159332P.
PR	19-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159330P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159331P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159637P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159638P.
PR	20-JUL-1999;	99US-0144352P.	PR	18-OCT-1999;	99US-0159584P.
PR	20-JUL-1999;	99US-0144632P.	PR	21-OCT-1999;	99US-0160741P.
PR	20-JUL-1999;	99US-0144884P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0160768P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160770P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160814P.
PR	21-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160815P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160816P.
PR	22-JUL-1999;	99US-0145089P.	PR	22-OCT-1999;	99US-0160980P.
PR	22-JUL-1999;	99US-0145192P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160989P.
PR	23-JUL-1999;	99US-0145224P.	PR	25-OCT-1999;	99US-0161404P.
PR	23-JUL-1999;	99US-0145226P.	PR	25-OCT-1999;	99US-0161405P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145913P.	PR	26-OCT-1999;	99US-0161359P.
PR	27-JUL-1999;	99US-0145918P.	PR	26-OCT-1999;	99US-0161360P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161361P.
PR	28-JUL-1999;	99US-0145951P.	PR	28-OCT-1999;	99US-0161920P.
PR	28-JUL-1999;	99US-0146386P.	PR	28-OCT-1999;	99US-0161992P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161993P.
PR	02-AUG-1999;	99US-0146389P.	PR	29-OCT-1999;	99US-0162142P.
PR	03-AUG-1999;	99US-0147038P.			
PR	04-AUG-1999;	99US-0147204P.			
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			

Query Match 51.9%; Score 42; DB 3; Length 772;  
Best Local Similarity 61.5%; Pred. No. 5.1e+02;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 2 LGNSSPRTQSPQN 14  
DB 667 LGNSAPISPPN 679

## RESULT 34

ADO61607  
ID ADO61607 standard; protein; 772 AA.  
XX  
AC ADO61607;  
XX  
DT 15-JUL-2004 (first entry)  
XX  
DE Transcription factor G324, SEQ ID 74.  
XX  
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;  
KW osmotic stress tolerance; cold tolerance; heat tolerance;  
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;  
KW glycosate resistance; flowering; fertility; seed development.  
XX  
OS Arabidopsis thaliana.  
XX  
FN WO2004031349-A2.  
XX  
PD 15-APR-2004.  
XX  
PF 18-SEP-2003; 2003WO-US030292.  
XX  
PR 18-SEP-2002; 2002US-0411837P.  
PR 17-DEC-2002; 2002US-0434166P.  
PR 24-APR-2003; 2003US-0465809P.  
XX  
PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
XX  
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;  
PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;  
PI  
DR WPI; 2004-330163/30.  
DR N-PSDB; ADO61606.  
XX  
XX New recombinant polynucleotide encoding transcription factor  
PT polypeptides, useful for producing transgenic plants with advantageous  
PT properties compared to a reference plant.  
XX  
PS Claim 16; SEQ ID NO 74; 510pp; English.  
XX  
XX The present invention relates to novel plant transcription factor  
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The  
CC sequences can be used to produce transgenic plants, which overexpress  
CC (II), where the transgenic plant has an altered trait as compared to a  
CC non-transgenic plant or wild-type plant. The transgenic plant comprises  
CC an altered trait selected from increased tolerance to abiotic stress,  
CC increased tolerance to osmotic stress, increased tolerance to cold,  
CC germination in heat, increased tolerance to freezing conditions,  
CC increased tolerance to low nitrogen conditions, increased tolerance to  
CC low phosphate conditions, increased tolerance to disease, including  
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,  
CC increased tolerance to multiple fungal pathogens, increased resistance to  
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,  
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance  
CC to sugars, altered carbon/nitrogen sensing, early flowering, late  
CC flowering, altered flower structure, loss of flower determinacy, reduced  
CC fertility, altered shoot meristem development, altered branching pattern,  
CC altered stem morphology, altered vascular tissue structure, reduced  
CC apical dominance, altered trichome density, altered trichome development,  
CC altered trichome structure, altered root development, altered shade  
CC avoidance, altered seed development, altered seed ripening, altered seed  
CC germination, slow growth, fast growth, altered cell differentiation,  
CC altered cell proliferation, altered cell expansion, altered phase change,  
CC altered senescence, abnormal embryo development, altered programmed cell  
CC death, lethality when overexpressed, altered necrosis patterns, increased  
CC plant size, increased biomass, large seedlings, dwarfed plants, dark  
CC green leaves, change in leaf shape, increased leaf size and mass, light  
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,  
CC altered seed coloration, altered seed size, altered seed shape, large

CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil  
CC content, altered seed protein content, altered seedprenyl content,  
CC altered leaf prenyl lipid content, increased anthocyanin levels, and  
CC decreased anthocyanin levels. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 772 AA;

Query Match 51.9%; Score 42; DB 8; Length 772;  
Best Local Similarity 61.5%; Pred. No. 5.1e+02;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LGNSPRTQSPQN 14  
| | | | : | | |  
Db 667 LGNSAPISPPN 679

## RESULT 35

ABR55649  
ID ABR55649 standard; peptide; 1224 AA.  
XX  
AC ABR55649;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Amino acid of an open reading frame of the pyranosone dehydratase gene.  
XX  
KW Pyranosone dehydratase; enzyme; microthecin; ascopyrone P; cortalcerone;  
KW Aphanomyces; 2-furyl-hydroxymethyl-ketone; 2-furylglyoxal;  
KW 4-deoxyglycero-hexo-2,3-dilucose; plant protectant; seed protectant;  
KW antimicrobial; antifungal.  
XX  
OS Phanerochaete chrysosporium.  
XX  
PN WO2003037918-A2.  
XX  
PD 08-MAY-2003.  
XX  
PF 30-OCT-2002; 2002WO-GB004916.  
XX  
PR 31-OCT-2001; 2001GB-00026164.  
PR 21-DEC-2001; 2001US-0343485P.  
XX  
PA (DANI-) DANISCO AS.  
XX  
PI Morgan AJ, Yu S, Weiergang I, Pedersen HC;  
XX  
DR WPI; 2003-441351/41.  
XX  
PT Novel polypeptide having pyranosone dehydratase activity, useful for  
PT preparing microthecin, ascopyrone P or cortalcerone which are useful as  
PT antimicrobial agents.  
XX  
PS Disclosure; Fig 5; 96pp; English.  
XX  
XX ABR55647-49 represent proteins encoded the open reading frames of the  
CC pyranosone dehydratase gene. This enzyme is useful for preparing  
CC microthecin, ascopyrone P (APP) or cortalcerone. Microthecin or  
CC cortalcerone, or its derivatives or isomers, are useful in preventing  
CC and/or inhibiting the growth of, and/or killing the pathogen Aphanomyces,  
CC preferably A. cohioides. The derivative of microthecin is 2-furyl-  
CC hydroxymethyl-ketone or 4-deoxy-glycero-hexo-2,3-dilucose. The derivative  
CC of cortalcerone is 2-furylglyoxal. Microthecin or cortalcerone, or its  
CC derivatives or isomers, are also useful for treating plant or plant seeds  
CC (such as sugar beet seeds) or as a plant or seed protectant. Microthecin  
CC and cortalcerone are useful as antimicrobial agents, particularly in food  
CC stuffs. Microthecin is useful as an antifungal for plant protection  
XX  
SQ Sequence 1224 AA;

Query Match 51.9%; Score 42; DB 6; Length 1224;

Best Local Similarity 50.0%; Pred. No. 8.3e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
Qy 2 LGNSPRTQSPQNC 15  
Db 67 IGLASPITETPQRC 80

RESULT 36  
AAO22566  
ID AAO22566 standard; protein; 2150 AA.

XX AAO22566;

XX 28-OCT-2002 (first entry)

XX Wooden leg (WOL) gene related protein SEQ ID No 17.

XX Wooden leg; WOL; vasculature; transgenic plant; agronomic; longer root;  
XX wood production; plant; promoter; tree; crop plant.

XX Arabidopsis thaliana.

XX WO200244337-A2.

XX 06-JUN-2002.

XX 29-NOV-2001; 2001WO-US045053.

XX 29-NOV-2000; 2000US-0253739P.

XX (UUNY ) UNIV NEW YORK STATE.

XX (HELA/) HELARIUTTA Y.

XX (MAHO/) MAHONEN A P.

XX (BONK/) BONKE A W M.

XX (KAUP/) KAUPPINEN L.

XX (RIIK/) RIIKONEN M.

XX Helariutta Y, Mahonen AP, Bonke AWM, Kauppinen L, Riikonen M;  
PI Benfey PN;  
XX WPI; 2002-599423/64.

XX Novel isolated polypeptide (WOODEN LEG) with ability to regulate a set of  
PT asymmetric cell divisions that establish vascular tissue in root and  
PT hypocotyl development, useful for improving agronomically valuable  
XX plants.

PS Example 1; Page 151-157; 187pp; English.  
XX The invention relates to an isolated WOODEN LEG (WOL) polypeptide,  
CC comprising 15 contiguous amino acids of a fully defined Arabidopsis  
CC WOODEN LEG protein sequence of 1057 amino acids as given in the  
CC specification, and to its encoding nucleic acid. The invention also  
CC relates to an amino acid sequence of domains of protein, e.g., N-terminal  
CC region, C-terminal domain, etc; or is a naturally occurring allelic  
CC variant of the above mentioned polypeptide sequence. Expression levels of  
CC the nucleic acid can be modified to improve the vasculature in transgenic  
CC plants and enhance the agronomic properties of such plants. Also the WOL  
CC promoter is used to drive expression of a heterologous coding sequence of  
CC trees to improve wood production. The WOL nucleic acid may be used as a  
CC molecular marker for a qualitative trait loci, e.g., longer roots or  
CC enhanced wood production, in molecular breeding of crop plants. The  
CC nucleic acid is also useful in DNA amplification assays to identify the  
CC endogenous WOL genes, WOL mutant alleles and/or WOL expression products  
CC in cultivars as compared to wild-type plants. They can also be used as  
CC markers for linkage analysis of qualitative trait loci. The WOL protein  
CC and/or antibodies can be used as diagnostic reagents in immunoassays to  
CC detect expression of the WOL gene in cultivars and wild-type plants. The  
CC WOL protein, its encoding nucleic acid, and its corresponding antibody  
CC are useful for improving agronomically valuable plants e.g., trees. This  
CC sequence represents a protein relating to the wooden leg (WOL) protein of  
CC the invention

XX SQ Sequence 2150 AA;  
Query Match 51.9%; Score 42; DB 5; Length 2150;  
Best Local Similarity 53.8%; Pred. No. 1.5e+03;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GNSPRTQSPQNC 15

Db 345 GNSPRTQSPQNC 357

RESULT 37

AAAR15781

ID AAR15781 standard; protein; 10 AA.

XX AAR15781;

XX 25-MAR-2003 (revised)

DT 09-JAN-2003 (revised)

DT 29-JAN-1992 (first entry)

XX Farnesyl-protein transferase inhibitor (34).

DE Farnesyl; transferase; FT; inhibitor; p21ras; rat.

XX Synthetic.

XX Key Location/Qualifiers

FT Peptide 7.10

XX WO9116340-A.

XX 31-OCT-1991.

XX 18-APR-1990; 90US-00510706.

XX 18-APR-1990; 90US-00510706.

XX 20-NOV-1990; 90US-00615715.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX Brown MS, Goldstein JL, Reiss Y;

XX WPI; 1991-339750/46.

XX Composn. comprising purified farnesyl-protein transferase - used to  
PT inhibit attachment of farnesyl moiety to RAS protein in malignant cells  
PT and to treat cancer.

XX Claim 25; Page 68; 87pp; English.

XX This peptide, the fragment indicated in the features, or the peptides  
CC represented in AAR15751-81, AAR14723 and AAR14711 inhibit the rat FTs  
CC represented in AAR14712-22. They show FT inhibition at an IC50 of 0.01-10  
CC microm. The most potent inhibitors are ones in which phenylalanine occurs  
CC at the third position of a tetrapeptide whose N-terminus is cysteine. The  
CC inhibitors have a farnesyl acceptor or inhibitor sequence within its  
CC structure and are capable of inhibiting the farnesylation of p21ras by  
CC FT. See also AAR14711-23 and AAQ14541-47. (Updated on 09-JAN-2003 to add  
CC missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 10 AA;

Query Match 50.6%; Score 41; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 6.9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TQSPQNC 15

Db 1 TQSPQNC 7

```

RESULT 38
AAR49742
ID AAR49742 standard; peptide; 10 AA.
XX AC AAR49742;
XX AC
XX DT 25-MAR-2003 (revised)
XX DT 08-AUG-1994. (first entry)
XX DE Farnesyltransferase-inhibitor.
XX DE
XX KW Farnesyltransferase-inhibitor; farnesyltransferase; FT; p21ras;
XX KW ras protein; farnesylation; cancer therapy.
XX OS Synthetic.
XX OS
XX PN WO9404561-A1.
XX PN
XX PD 03-MAR-1994.
XX PF 24-AUG-1993; 93WO-US008062.
XX PR 24-AUG-1992; 92US-00935087.
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX PA (GETH ) GENENTECH INC.
XX PI Brown MS, Goldstein JL, Reiss Y, Marsters JC;
XX WPI; 1994-083105/10.
XX DR
XX PT New farnesyl-transferase inhibitors - used for inhibiting attachment of a
XX PT farnesyl moiety to a p13ras protein in malignant cells.
XX PS Disclosure; Page 33; 183pp; English.
XX CC
XX CC Peptides given in AAR49741-75, AAR49777-78 and AAR49785-88, which include
XX CC a family of tetrapeptides based on the recognition site (AAR49776) of
XX CC farnesyltransferase (FT), are potential anticancer agents that inhibit
XX CC FT, thereby preventing expression of p21ras. (Updated on 25-MAR-2003 to
XX CC correct PN field.)
XX SQ Sequence 10 AA;

Query Match 50.6%; Score 41; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TQSPQNC 15
Db |||||
1 TQSPQNC 7

RESULT 40
AAR04434
ID AAR04434 standard; peptide; 10 AA.
XX AC AAR04434;
XX AC
XX DT 30-JUL-1997 (first entry)
XX DE
XX DE Farnesyl transferase peptide inhibitor used in cancer treatment.
XX KW Farnesyl transferase; inhibitor; cancer; tumour; neoplasia; prenyl;
XX KW ras protein; K-ras B; malignant; detection; identification.
XX OS Synthetic.
XX OS
XX PN WO9634113-A2.
XX PN
XX PD 31-OCT-1996.
XX PF 29-APR-1996; 96WO-US005969.
XX PR 27-APR-1995; 95US-00429964.
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX PA
XX PI Brown MS, Goldstein JL, James GL;
XX PI WPI; 1996-497642/49.
XX DR
XX DR Assay for farnesyl transferase activity - by determining ability to
XX PT transfer farnesyl moiety to K-Ras B protein, partic. useful for
XX PT identifying inhibitors.
XX PS Disclosure; Page 33; 257pp; English.

RESULT 39
AAR77804
ID AAR77804 standard; protein; 10 AA.
XX AC AAR77804;
XX AC
XX DT 25-MAR-2003 (revised)
XX DT 23-JAN-1996 (first entry)
XX DE Farnesyl transferase inhibitor peptide, TQSPQNC5IM.
XX DE
XX KW Farnesyl transferase; inhibitor; cancer; ras; p21.
XX KW
XX OS Synthetic.
XX OS
XX PN US5420245-A.
XX PN
XX PD 30-MAY-1995.
XX PD
XX PF 03-APR-1992; 92US-00863169.

```

```

XX 18-APR-1990; 90US-00510706.
XX PR 20-NOV-1990; 90US-00615715.
XX PR 16-JAN-1992; 92US-00822011.
XX PA (TEXA ) UNIV TEXAS.
XX PI Reiss Y, Goldstein JL, Brown MS;
XX WPI; 1995-206308/27.
XX DE
XX DE New farnesyl transferase inhibitor peptide(s) - based on farnesyl
XX PT acceptor substrate carboxy terminal sequences, used for the treatment of
XX PT cancer.
XX PS Claim 2; Col 61; 55pp; English.
XX CC
XX CC AAR77801-R77804 are peptide inhibitors of farnesyl transferase. They all
XX CC have a carboxyl terminal sequence obeying the generic formula -CRAAX,
XX CC where C= cysteine, A= any aliphatic, aromatic or hydroxy amino acid and
XX CC X= any normal amino acid. Farnesyl transferase is involved in the
XX CC farnesylation of various cellular proteins including the cancer related
XX CC ras proteins. The transforming activity of ras is dependent on the
XX CC localisation of the protein to membranes, a property which is thought to
XX CC be dependent upon the addition of farnesyl groups. The peptide inhibitors
XX CC are useful for treating cancers and ras-related cancers in particular.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX SQ Sequence 10 AA;

Query Match 50.6%; Score 41; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TQSPQNC 15
Db |||||
1 TQSPQNC 7

RESULT 40
AAR04434
ID AAR04434 standard; peptide; 10 AA.
XX AC AAR04434;
XX AC
XX DT 30-JUL-1997 (first entry)
XX DE
XX DE Farnesyl transferase peptide inhibitor used in cancer treatment.
XX KW Farnesyl transferase; inhibitor; cancer; tumour; neoplasia; prenyl;
XX KW ras protein; K-ras B; malignant; detection; identification.
XX OS Synthetic.
XX OS
XX PN WO9634113-A2.
XX PN
XX PD 31-OCT-1996.
XX PF 29-APR-1996; 96WO-US005969.
XX PR 27-APR-1995; 95US-00429964.
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX PA
XX PI Brown MS, Goldstein JL, James GL;
XX PI WPI; 1996-497642/49.
XX DR
XX DR Assay for farnesyl transferase activity - by determining ability to
XX PT transfer farnesyl moiety to K-Ras B protein, partic. useful for
XX PT identifying inhibitors.
XX PS Disclosure; Page 33; 257pp; English.

```

XX AA04433-W04465 are peptide inhibitors of farnesyl transferase (FT)  
 CC activity. The peptides block the attachment of prenyl groups to ras  
 CC proteins in malignant cells of patients suffering from cancer or a  
 CC precancerous state and as such are used to treat cancer. The peptides  
 CC were identified by determining the ability of candidate substances to  
 CC inhibit a FT enzyme, by inhibiting the transfer of a farnesyl moiety to a  
 CC K-RasB protein  
 XX  
 SQ Sequence 10 AA;  
 Query Match 50.6%; Score 41; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 9 TQSPQNC 15  
 Db 1 TQSPQNC 7  
 RESULT 41  
 AAU53546  
 ID AAU53546 standard; protein; 68 AA.  
 XX  
 AC AAU53546;  
 XX  
 DT 27-FEB-2002 (first entry)  
 XX  
 DE Propionibacterium acnes immunogenic protein #14442.  
 XX  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 FN WO200181581-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 20-APR-2001; 2001WO-US012865.  
 XX  
 PR 21-APR-2000; 2000US-0199047P.  
 PR 02-JUN-2000; 2000US-0208841P.  
 PR 07-JUL-2000; 2000US-0216747P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 WPI; 2001-616774/71.  
 DR N-PSDB; AAS59560.  
 XX  
 PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris.  
 XX  
 PS Example 1; SEQ ID NO 14741; 1069pp; English.  
 XX  
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 68 AA;  
 Query Match 50.6%; Score 41; DB 4; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 54;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 SSPRTQSP 12  
 Db 2 SSPRTQSP 9  
 RESULT 42  
 ABM50065  
 ID ABM50065 standard; protein; 68 AA.  
 XX  
 AC ABM50065;  
 XX  
 DT 20-OCT-2003 (first entry)  
 XX  
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #14741.  
 XX  
 KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
 KW immunostimulant; immune response; vaccine.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 FN WO2003033515-A1.  
 XX  
 PD 24-APR-2003.  
 XX  
 PF 11-OCT-2002; 2002WO-US032727.  
 XX  
 PR 15-OCT-2001; 2001US-00978825.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
 PI Barth B, Vallie-Douglas J;  
 XX  
 WPI; 2003-381789/36.  
 DR N-PSDB; ACF64489.  
 XX  
 PT New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a P. acnes protein.  
 XX  
 PS Example 1; SEQ ID NO 14741; 1481pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
 CC encoding a Propionibacterium acnes protein. The invention also relates to  
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
 CC immunogenic fragments of P. acnes polypeptides. The invention  
 CC additionally encompasses expression vectors and host cells comprising a  
 CC polynucleotide of the invention; antibodies against polypeptides of the  
 CC invention; fusion proteins comprising a polypeptide of the invention; a  
 CC method for stimulating an immune response specific for a P. acnes  
 CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,  
 CC via this method; a vaccine composition comprising P. acnes polypeptides,  
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
 CC antigen-presenting cells that express the polypeptide; a method and kit  
 CC for detecting or determining the presence or absence of P. acnes in a  
 CC patient; and a method for inhibiting the development of P. acnes in a  
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

CC proteins, T cell populations or antigen-presenting cells that express the  
 CC polypeptides are useful for diagnosing, preventing or treating acne  
 CC vulgaris, or for stimulating an immune response specific for a P. acnes  
 CC protein. The polynucleotides can also be used as probes or primers for  
 CC nucleic acid hybridisation. The vaccine composition is useful for the  
 CC stimulation of an immune response against P. acnes, or for treating acne,  
 CC and the kit is useful for performing a diagnostic assay. The present  
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
 CC reading frame) contained within the P. acnes polynucleotides of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX

SQ Sequence 68 AA;

Query Match 50.6%; Score 41; DB 6; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 54;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 SSPRTQSP 12  
 |||||  
 Db 2 SSPRTQSP 9

RESULT 43  
 ADL06330  
 ID ADL06330 standard; protein; 539 AA.

AC ADL06330;

DT 06-MAY-2004 (first entry)

DE Amorpha fruticosa 4-coumaric acid:co-enzyme A ligase.

KW 4-coumaric acid:co-enzyme A ligase; CoA; 4CL; transgenic; plant; enzyme.

XX Amorpha fruticosa.

FT Key Location/Qualifiers

FT Misc-difference 465..466

FT /note= "Encoded by GTTCCAATG"

FN CN1390940-A.

PD 15-JAN-2003.

PF 19-JUL-2002; 2002CN-00132631.

PR 19-JUL-2002; 2002CN-00132631.

PA (UYDA-) UNIV DALIAN SCI & ENG.

PI An L, Liu W, Su Q;

XX WPI; 2003-469215/45.

DR N-PSDB; ADL06329.

XX 4-coumaric acid: CoA ligase gene and its clone.

PT Disclosure; Page 7-8; 11pp; Chinese.

XX The present invention relates to the isolation of a plant 4-coumaric  
 CC acid:Co-enzyme A (CoA) ligase (4CL) gene from Amorpha fruticosa, and the  
 CC protein it encodes. Also disclosed is a clonal carrier pUC19/4CL  
 CC containing 4CL, and an Escherichia coli cell JW109/pUC19/4CL containing  
 CC the carrier. It can be used to configure genetically engineered bacteria  
 CC for transgenic plants. The present sequence represents Amorpha fruticosa  
 CC 4CL protein.

SQ Sequence 539 AA;

Query Match 50.6%; Score 41; DB 7; Length 539;  
 Best Local Similarity 61.5%; Pred. No. 5e+02;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPNC 15  
 |||||  
 Db 374 GNSLPRNQSGEIC 386

RESULT 44  
 AAW26130  
 ID AAW26130 standard; protein; 542 AA.

XX AAW26130;

DT 17-OCT-2003 (revised)

DT 21-NOV-1997 (first entry)

XX 4-coumaric acid coenzyme A ligase.

XX 4-coumaric acid coenzyme A ligase; 4CL gene; tobacco; Nicotiana tabacum;  
 KW lignin.

XX Nicotiana tabacum; L. SR1 strain.

XX JP09173069-A.

PD 08-JUL-1997.

PF 22-DEC-1995; 95JP-00334834.

PR 22-DEC-1995; 95JP-00334834.

XX (MITY) MITSUBISHI PAPER MILLS LTD.

XX WPI; 1997-397027/37.

DR N-PSDB; AAT80088.

XX 4-coumaric acid:coenzyme A ligase gene - used to reduce the lignin  
 PT content of plants.

PS Claim 9; Page 6-8; 10pp; Japanese.

XX This sequence represents the 4-coumaric acid coenzyme A ligase (4CL) from  
 CC Nicotiana tabacum strain SR1. The DNA encoding this sequence, or the 4CL  
 CC gene fragments shown in AAT80086 and AAT80087, are introduced into a  
 CC plant in the method of the invention. The method of the invention is for  
 CC the reduction of lignin in a plant. The lignin content in a transformed  
 CC plant can be reduced using this method. (Updated on 17-OCT-2003 to  
 CC standardise OS field)

SQ Sequence 542 AA;

Query Match 50.6%; Score 41; DB 2; Length 542;

Best Local Similarity 61.5%; Pred. No. 5.1e+02;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPNC 15  
 |||||  
 Db 377 GNSLPRNQSGEIC 389

RESULT 45

ADR86215

ID ADR86215 standard; protein; 747 AA.

XX ADR86215;

DT 04-NOV-2004 (first entry)

XX Aspergillus fumigatus essential gene protein #265.

XX Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;  
 KW drug screening.

```
OS Aspergillus fumigatus.
XX WO2004067709-A2.
XX 12-AUG-2004.
XX 16-JAN-2004; 2004WO-US001099.
XX 17-JAN-2003; 2003US-0441281P.
XX 13-JUN-2003; 2003US-0478196P.
XX (ELIT-) ELITRA PHARM INC.
XX (ELIT-) ELITRA CANADA LTD.
XX Jiang B, Hu W, Lemieux S, Roemer T;
XX WPI; 2004-594200/57.
XX N-PSDB; ADR85628.
XX New purified or isolated Aspergillus fumigatus nucleic acid molecule
XX encoding a gene product, useful for diagnosing and/or treating invasive
XX fungal infections, such as Farmer's lung disease.
XX Claim 1; SEQ ID NO 3265; 164pp; English.
XX The present invention relates to Aspergillus fumigatus genes that are
XX essential and are potential targets for drug screening. The methods and
XX compositions of the present invention are useful for diagnosing and/or
XX treating invasive Aspergillus fumigatus infection, including the allergic
XX forms of the disease, such as Farmer's lung disease. They can also be
XX used in various drug discovery purposes, such as expression of the
XX recombinant protein, hybridization assay and construction of nucleic acid
XX arrays. The present sequence represents an Aspergillus fumigatus
XX essential gene protein sequence, used during diagnosis and drug
XX development in the invention. These genes share a high degree of sequence
XX conservation with known essential genes of candida albicans. The sequence
XX data for this patent is not represented in the printed specification, but
XX was obtained in electronic format from WIPO.
XX Sequence 747 AA;
XX
XX Query Match 50.6%; Score 41; DB 8; Length 747;
XX Best Local Similarity 69.2%; Pred. No. 7.1e+02;
XX Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX Qy 1 LLGNSSPRTQSPQ 13
XX | | | | |
XX 204 LFTNSSPPEQSPQ 216
XX
XX RESULT 46
XX ADQ66060
XX ID ADQ66060 standard; protein; 791 AA.
XX AC ADQ66060;
XX 07-OCT-2004 (first entry)
XX Novel human protein sequence #1033.
XX osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
XX gene therapy; diagnostic marker; morbid state; osteoporosis;
XX neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
XX cancer.
XX Homo sapiens.
XX EP1440981-A2.
XX 28-JUL-2004.
XX 21-JAN-2004; 2004EP-00001196.
XX
XX PR 21-JAN-2003; 2003JP-00102206.
XX 09-MAY-2003; 2003JP-00131392.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Nagai K, Irie R;
XX WPI; 2004-535376/52.
XX N-PSDB; ADQ63872.
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 3221; 2449pp; English.
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
XX polypeptides, sequences hybridizing to these nucleotides, sequences
XX encoding partial polypeptides and sequences having 70% or 90% identity to
XX the nucleotide and protein sequences. The nucleotides and polypeptides
XX are useful as diagnostic markers or therapeutic target for the diseases
XX or morbid states. They are also useful for treating osteoporosis,
XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,
XX dementia and various cancers. This sequence corresponds to a protein
XX sequence of the invention.
XX Sequence 791 AA;
XX
XX Query Match 50.6%; Score 41; DB 8; Length 791;
XX Best Local Similarity 64.3%; Pred. No. 7.6e+02;
XX Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX Qy 1 LLGNSSPRTQSPQ 14
XX | | | | |
XX 29 LLGSKRPQSYQS 42
XX
XX RESULT 47
XX ADK52121
XX ID ADK52121 standard; protein; 871 AA.
XX AC ADK52121;
XX 20-MAY-2004 (first entry)
XX Human atopic dermatitis/psoriasis-associated protein #36.
XX Human atopic dermatitis; psoriasis; dermatological; anti-inflammatory;
XX antipsoriatic; rash.
XX Homo sapiens.
XX WO2004016785-A1.
XX 26-FEB-2004.
XX 06-AUG-2003; 2003WO-JP0099999.
XX 06-AUG-2002; 2002JP-00229319.
XX 14-MAY-2003; 2003JP-00136544.
XX (GENO-) GENOX RES INC.
XX (UKJU-) UNIV JUNTENDO.
XX Itoh M, Ogawa K, Shinagawa A, Sudo H, Ogawa H, Ra C;
XX Mitsuishi K;
XX WPI; 2004-214514/20.
XX N-PSDB; ADK52035.
XX
XX Detecting atopic dermatitis or psoriasis comprises assaying levels of
XX expression of an indicator gene at a rash site and non-rash site of a
XX person with atopic dermatitis or psoriasis.
```

XX Example 2; SEQ ID NO 154; 484pp; Japanese.

XX The invention relates to detecting atopic dermatitis or psoriasis

XX comprising assaying the levels of expression of an indicator gene at a

XX rash site and non-rash site of a person with atopic dermatitis or

XX psoriasis, comparing these levels with those of a healthy person, and

XX determining that if the levels of indicators are higher or lower, then

XX this indicates the disease. Also included are a reagent for detecting

XX atopic dermatitis or psoriasis, a kit for screening for treatments, a

XX transgenic non human vertebrate animal models for the diseases, an agent

XX for inducing the diseases in mice and a DNA chip for assaying for the

XX indicator genes. The method is used for treatment, detection and animal

XX models for research of atopic dermatitis and psoriasis. The present

XX sequence is a protein encoded by an indicator gene of the invention.

XX SQ Sequence 871 AA;

Query Match 50.6%; Score 41; DB 8; Length 871;

Best Local Similarity 64.3%; Pred. No. 8.4e+02;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQN 14

Db 29 LLGRSKPRPQSQS 42

|||||

|||||

RESULT 48

ADNR09093

ID ADR09093 standard; protein; 871 AA.

XX AC ADR09093;

XX DT 04-NOV-2004 (first entry)

XX DE Human protein useful for treating neurological disease Seq 2599.

XX KW human; oligo-capping method; diagnostic marker; gene therapy;

XX KW osteoporosis; neurological disease; Alzheimer's disease;

XX KW Parkinson's disease; dementia; short memory; cancer;

XX KW sense or motor function; emotional reaction; fear response; panic;

XX KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;

XX KW tranquiliser.

XX OS Homo sapiens.

XX PN EP1447413-A2.

XX PD 18-AUG-2004.

XX PF 12-FEB-2004; 2004EP-00003145.

XX PR 14-FEB-2003; 2003JP-00102207.

XX PR 09-MAY-2003; 2003JP-00131452.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;

XX Wakamatsu A, Ishii S, Nagai K, Irie R;

XX WPI; 2004-583265/57.

XX DR N-PSDB; ADR07137.

XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,

XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 2599; 2686pp; English.

XX This invention relates to novel, isolated full length human cDNA

XX molecules and the encoded proteins thereof. Specifically, it refers to

XX cDNA clones obtained by an oligo-capping method, where none of these

XX clones are identical to any known human mRNAs. The present invention

XX describes an immunoassay to identify agonists and antagonists, as well as

CC antibodies, antisense molecules and siRNAs that can all be used to bind

CC to and modulate expression of the cDNA molecules. As such, these

CC molecules are useful for diagnostic markers or therapeutic targets for

CC the various diseases or morbid states. In particular, they are useful in

CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's

CC disease, Parkinson's disease, dementia, short memory and various cancers,

CC as well as for maintaining equilibrium of sense or motor function, and

CC for treating emotional reaction, fear response and panic. Accordingly,

CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,

CC cyostatic and tranquiliser activities. This polypeptide is a protein

CC encoded by a full length human cDNA sequence of the invention. NOTE: This

CC sequence is not given in the sequence listing of the specification but

CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-

CC office.

XX SQ Sequence 871 AA;

Query Match 50.6%; Score 41; DB 8; Length 871;

Best Local Similarity 64.3%; Pred. No. 8.4e+02;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQN 14

Db 29 LLGRSKPRPQSQS 42

|||||

|||||

RESULT 49

ABBS58985

ID ABBS58985 standard; protein; 1937 AA.

XX AC ABBS58985;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 3747.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX DR N-PSDB; ABL03088.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signaling and cell-cell

XX interactions.

XX Disclosure; SEQ ID NO 3747; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-

XX ABBS72072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at fip.wipo.int/pub/published\_pct\_sequences



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: August 25, 2005, 11:33:48 ; Search time 168 Seconds  
(without alignments)  
45.721 Million cell updates/sec

Title: US-10-803-541-2

Perfect score: 81

Sequence: 1 LLGNSSPRTQSPQNC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	664	1 LAMA_HUMAN	P02545 homo sapien
2	73	90.1	665	1 LAMA_RAT	P48679 rattus norv
3	69	85.2	135	2 Q8ENX5	Q8bnx5 mus musculu
4	69	85.2	665	1 LAMA_MOUSE	P48678 mus musculu
5	69	85.2	665	2 Q91WF2	Q91wf2 mus musculu
6	69	85.2	665	2 Q9DC21	Q9dc21 mus musculu
7	49	60.5	886	2 Q6A050	Q6a050 mus musculu
8	48	59.3	172	2 Q6ZRC6	Q6zrc6 homo sapien
9	48	59.3	852	2 Q6CG31	Q6cg31 yarrowia li
10	46	56.8	514	1 HUTH_PSEBFL	Q8vmr3 pseudomonas
11	45	55.6	162	2 Q9FWN9	Q9ftw9 oryza sativ
12	44	54.3	180	2 Q8BPZ3	Q8bpz3 mus musculu
13	44	54.3	286	2 Q96MH2	Q96mh2 homo sapien
14	44	54.3	443	2 Q9VLQ0	Q9vlq0 drosophila
15	44	54.3	635	2 P91092	P91092 caenorhabdi
16	44	54.3	643	2 Q95Q79	Q95q79 caenorhabdi
17	44	54.3	826	2 Q8H852	Q8h852 oryza sativ
18	44	54.3	970	1 PSUL_YEAST	P53550 saccharomyc
19	43	53.1	559	2 Q6CT64	Q6ct64 yarrowia li
20	43	53.1	2001	2 Q8RZN9	Q8rzn9 oryza sativ
21	42	51.9	161	2 Q9DSW7	Q9dsw7 ascovirus d
22	42	51.9	191	2 Q38966	Q38966 arabidopsis
23	42	51.9	291	2 Q7D0I6	Q7d0i6 agrobacteri
24	42	51.9	690	1 PERO_DROME	Q01603 drosophila
25	42	51.9	772	2 Q9LSE3	Q9lse3 arabidopsis
26	42	51.9	901	2 Q6BY30	Q6by30 debaryomyc
27	42	51.9	962	1 SUUR_DROER	P59597 drosophila
28	42	51.9	2150	2 Q23863	Q23863 dictyosteli
29	42	51.9	4708	2 Q7TPH6	Q7tph6 mus musculu
30	41	50.6	148	2 Q8YL47	Q8yl47 anabaena sp
31	41	50.6	158	2 Q7R2Z1	Q7r2z1 giardia lam

## ALIGNMENTS

32	41	50.6	240	1	RCF2_PSEAE	Q86638 pseudomonas
33	41	50.6	279	2	Q869Y4	Q869y4 dictyosteli
34	41	50.6	439	2	Q8YUK2	Q8yuk2 anabaena sp
35	41	50.6	446	2	Q68CL1	Q68cl1 homo sapien
36	41	50.6	489	1	MURE_COXBU	Q83f28 coxiella bu
37	41	50.6	540	2	Q8W558	Q8w558 amorpha fru
38	41	50.6	542	1	4CL2_TOBAC	Q24146 nicotiana t
39	41	50.6	542	2	Q42943	Q42943 nicotiana t
40	41	50.6	556	2	Q6TUF8	Q6tuf8 rattus norv
41	41	50.6	577	2	Q8PIR2	Q8pir2 xanthomonas
42	41	50.6	593	1	CDYL_MOUSE	Q9wtk2 mus musculu
43	41	50.6	599	2	Q64033	Q64033 mus sp. ant
44	41	50.6	871	2	Q96DR7	Q96dr7 homo sapien
45	41	50.6	871	2	Q6AZ96	Q6az96 homo sapien
46	41	50.6	871	2	Q6Q8Q8	Q6q8q8 homo sapien
47	41	50.6	1014	2	Q6DRCS	Q6drc5 brachydanio
48	41	50.6	2038	1	FSH_DROME	P13709 drosophila
49	41	50.6	2038	2	Q9W3L3	Q9w3l3 drosophila
50	40.5	50.0	470	1	SA12_XENLA	P40649 xenopus lae
51	40.5	50.0	567	2	Q6NUE6	Q6nue6 xenopus lae
52	40	49.4	54	2	Q82RX5	Q82rx5 streptomyce
53	40	49.4	94	2	Q9PA03	Q9pa03 xylella fas
54	40	49.4	170	2	Q8N1L1	Q8n1l1 homo sapien
55	40	49.4	199	1	SOXE_SULAC	Q53765 sulfolobus
56	40	49.4	212	2	Q97WT6	Q97wt6 sulfolobus
57	40	49.4	257	1	HMMA_BRARE	Q03357 brachydanio
58	40	49.4	294	2	Q73708	Q73708 brachydanio
59	40	49.4	294	2	Q6PCA5	Q6pca5 brachydanio
60	40	49.4	302	2	Q9VIT1	Q9vit1 drosophila
61	40	49.4	311	2	Q8INV9	Q8inv9 drosophila
62	40	49.4	311	2	Q8ENQ9	Q8enq9 drosophila
63	40	49.4	318	2	Q8A804	Q8a804 bacteroides
64	40	49.4	345	2	Q8DH97	Q8dh97 synecococc
65	40	49.4	370	2	Q6CNE8	Q6cne8 kluyveromyc
66	40	49.4	436	2	Q63DS5	Q63ds5 bacillus ce
67	40	49.4	452	2	Q8INW0	Q8inw0 drosophila
68	40	49.4	454	2	Q6LDB6	Q6ldb6 rattus sp.
69	40	49.4	462	2	Q8IZ65	Q8iz65 homo sapien
70	40	49.4	462	2	Q9H8Y6	Q9h8y6 homo sapien
71	40	49.4	462	2	Q9HAJ1	Q9haj1 homo sapien
72	40	49.4	551	2	Q8UAQ1	Q8uaq1 agrobacteri
73	40	49.4	622	2	Q9EQS1	Q9eqs1 cricetulus
74	40	49.4	641	2	Q6BY25	Q6by25 debaryomyc
75	40	49.4	690	2	Q7CS58	Q7cs58 agrobacteri
76	40	49.4	750	2	Q7Z6G6	Q7z6g6 homo sapien
77	40	49.4	758	1	MEFE_PHOLL	Q7mz74 photorhabd
78	40	49.4	758	2	Q9GQL5	Q9gql5 trypanosoma
79	40	49.4	768	2	Q6IVB5	Q6ivb5 homo sapien
80	40	49.4	828	2	Q6FP48	Q6fp48 candida gla
81	40	49.4	870	2	Q9XTA4	Q9xta4 bos taurus
82	40	49.4	934	2	Q6FY14	Q6fy14 candida gla
83	40	49.4	950	2	Q6AX98	Q6ax98 xenopus lae
84	40	49.4	955	2	Q9Y5K9	Q9y5k9 homo sapien
85	40	49.4	962	2	Q80U28	Q80u28 mus musculu
86	40	49.4	962	2	Q91YY9	Q91yy9 mus musculu
87	40	49.4	1004	2	Q658V6	Q658v6 homo sapien
88	40	49.4	1008	2	Q8AY57	Q8ay57 fundulus he
89	40	49.4	1012	1	IF2C_PHAVU	P57997 phaseolus v
90	40	49.4	1017	1	AI42_CHICK	P24797 gallus gall
91	40	49.4	1017	2	Q90X34	Q90x34 brachydanio
92	40	49.4	1017	2	Q9DGL5	Q9dgl5 brachydanio
93	40	49.4	1020	1	AI42_HUMAN	P50993 homo sapien
94	40	49.4	1020	1	AI42_RAT	P06686 rattus norv
95	40	49.4	1020	2	Q6PIE5	Q6pie5 mus musculu
96	40	49.4	1022	2	Q6ZPAG0	Q6zpag0 xenopus lae
97	40	49.4	1022	2	Q6ZQ49	Q6zq49 mus musculu
98	40	49.4	1025	2	Q91WH7	Q91wh7 mus musculu
99	40	49.4	1030	1	ATHA_XENLA	Q92126 xenopus lae
100	40	49.4	1032	1	ATHA_MOUSE	Q64436 mus musculu

RESULT 1  
LAMA\_HUMAN STANDARD; PRT; 664 AA.  
ID LAMA\_HUMAN  
AC P02545; P02546; Q96918; Q96J24;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 20-MAR-1987 (Rel. 04, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Lamin A/C (70 kDa lamin)  
GN Name=LMA; Synonyms=LMN1;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS A AND C).  
RX MEDLINE=86118697; PubMed=3453101;  
RA McKeon F.D., Kirschner M.W., Caput D.;  
RT "Homologies in both primary and secondary structure between nuclear  
RT envelope and intermediate filament proteins.";  
RL Nature 319:463-468(1986).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS A AND C), AND SEQUENCE OF 583-644.  
RX MEDLINE=86313596; PubMed=3462705;  
RA Fisher D.Z., Chaudhary N., Blobel G.;  
RT "cDNA sequencing of nuclear lamins A and C reveals primary and  
RT secondary structural homology to intermediate filament proteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:6450-6454(1986).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM C).  
RC TISSUE=Kidney, and Lung;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Faney J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.B.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 375-664 FROM N.A. (ISOFORM ADELTA10).  
RC TISSUE=Colon;  
RX MEDLINE=96199170; PubMed=8621584; DOI=10.1074/jbc.271.16.9249;  
RA Machiels B.M., Zorenc A.H., Endert J.M., Kuipers H.J., van Eys G.J.,  
RA Ramaekers F.C., Broers J.L.;  
RT "An alternative splicing product of the lamin A/C gene lacks exon  
RT 10.";  
RL J. Biol. Chem. 271:9249-9253(1996).  
RN [5]  
RP INTERACTIONS WITH TMPO-ALPHA AND RB1.  
RX PubMed=12475961; DOI=10.1093/mbc.E02-07-0450;  
RA Markiewicz E., Dechat T., Foissner R., Quinlan R.A., Hutchison C.J.;  
RT "Lamin A/C binding protein Larp2alpha is required for nuclear anchorage  
RT of retinoblastoma protein.";  
RL Mol. Biol. Cell 13:4401-4413(2002).  
RN [6]  
RP VARIANTS EDM2 TRP-453; PRO-527 AND PRO-530.  
RX MEDLINE=99178265; PubMed=10080180; DOI=10.1038/6799;  
RA Bonne G., Di Barletta M.R., Varnous S., Becane H.-M., Hammouda E.-H.,  
RA Merlini L., Muntoni F., Greenberg C.R., Gary F., Urtizberea J.-A.,

RA Duboc D., Fardeau M., Toniolo D., Schwartz K.;  
RT "Mutations in the gene encoding lamin A/C cause autosomal dominant  
RT Emery-Dreifuss muscular dystrophy.";  
RL Nat. Genet. 21:285-288(1999).  
RN [7]  
RP VARIANTS CMD1A GLY-60; ARG-85; LYS-195 AND GLY-203.  
RX MEDLINE=20037766; PubMed=10580070; DOI=10.1056/NEJM199912023412302;  
RA Fackin D., Macrae C., Sasaki T., Wolff M.R., Porcu M., Freneaux M.,  
RA Atherton J., Vidallet H.J. Jr., Spudich S., De Girolami U.,  
RA Seidman J.G., Seidman C.E.;  
RT "Missense mutations in the rod domain of the lamin A/C gene as causes  
RT of dilated cardiomyopathy and conduction-system disease.";  
RL N. Engl. J. Med. 341:1715-1724(1999).  
RN [8]  
RP VARIANT EDM3 TYR-222.  
RX MEDLINE=20206333; PubMed=10739764;  
RA Raffaele di Barletta M., Ricci E., Galluzzi G., Tonali P., Mora M.,  
RA Morandi L., Romorini A., Voit T., Orstavik K.H., Merlini L.,  
RA Trevisan C., Biancalana V., Housmanowa-Petrusewicz I., Bione S.,  
RA Ricotti R., Schwartz K., Bonne G., Toniolo D.;  
RT "Different mutations in the LMNA gene cause autosomal dominant and  
RT autosomal recessive Emery-Dreifuss muscular dystrophy.";  
RL Am. J. Hum. Genet. 66:1407-1412(2000).  
RN [9]  
RP VARIANT PFLD GLN-482.  
RX MEDLINE=20056129; PubMed=10587585; DOI=10.1093/hmg/9.1.1.109;  
RA Cao H., Hegeler R.A.;  
RT "Nuclear lamin A/C R482Q mutation in Canadian kindreds with Dunnigan-  
RT type familial partial lipodystrophy.";  
RL Hum. Mol. Genet. 9:109-112(2000).  
RN [10]  
RP VARIANTS PFLD LEU-482 AND TRP-482.  
RX MEDLINE=20120714; PubMed=10655060; DOI=10.1038/72807;  
RA Shackleton S., Lloyd D.J., Jackson S.N.J., Evans R., Niermeijer M.F.,  
RA Singh B.M., Schmidt H., Brabant G., Kumar S., Durrington P.N.,  
RA Gregory S., O'Rahilly S., Trembath R.C.;  
RT "LMNA encoding lamin A/C is mutated in partial lipodystrophy.";  
RL Nat. Genet. 24:153-156(2000).  
RN [11]  
RP VARIANTS PFLD ASP-465; GLN-482; TRP-482 AND HIS-582.  
RX MEDLINE=20206314; PubMed=10739751;  
RA Speckman R.A., Garg A., Du F., Bennett L., Veille R., Ariloglu E.,  
RA Taylor S.I., Lovett M., Bowcock A.M.;  
RT "Mutational and haplotype analyses of families with familial partial  
RT lipodystrophy (Dunnigan variety) reveal recurrent missense mutations  
RT in the globular C-terminal domain of lamin A/C.";  
RL Am. J. Hum. Genet. 66:1192-1198(2000).  
RN [12]  
RP VARIANT EDM2 PRO-133.  
RX PubMed=11503164; DOI=10.1002/ajmg.1463;  
RA Brown C.A., Lanning R.W., McKinney K.Q., Salvino A.R., Cherniske E.,  
RA Crowe C.A., Darras B.T., Gominak S., Greenberg C.R., Grossmann C.,  
RA Heydemann P., Mendell J.R., Pober B.R., Sasaki T., Shapiro F.,  
RA Simpson D.A., Suchowersky O., Spence J.E.;  
RT "Novel and recurrent mutations in lamin A/C in patients with Emery-  
RT Dreifuss muscular dystrophy.";  
RL Am. J. Med. Genet. 102:359-367(2001).  
RN [13]  
RP VARIANT CMT2B1 CYS-298.  
RX PubMed=11799477;  
RA De Sandre-Giovannoli A., Chaoch M., Kozlov S., Vallat J.-M.,  
RA Tazir M., Kassouri N., Szepietowski P., Hammadouché T.,  
RA Vandenberghé A., Stewart C.L., Grid D., Levy N.;  
RT "Homozygous defects in LMNA, encoding lamin A/C nuclear-envelope  
RT proteins, cause autosomal recessive axonal neuropathy in human  
RT (Charcot-Marie-Tooth disorder type 2) and mouse.";  
RL Am. J. Hum. Genet. 70:726-736(2002).  
RN [14]  
RP ERRATUM.  
RA De Sandre-Giovannoli A., Chaoch M., Kozlov S., Vallat J.-M.,  
RA Tazir M., Kassouri N., Szepietowski P., Hammadouché T.,  
RA Vandenberghé A., Stewart C.L., Grid D., Levy N.;  
RL Am. J. Hum. Genet. 70:1075-1075(2002).

[15]  
 RP VARIANT WADA HIS-527.  
 RX PubMed=12075506;  
 RA Novelli G., Muchir A., Sangiulio F., Helbling-Leclerc A.,  
 RA D'Apice M.R., Massart C., Capon P., Sbraccia P., Federici M.,  
 RA Lauro R., Tufisco C., Pallotta R., Scarano G., Dallapiccola B.,  
 RA Merlini L., Bonne G.;  
 RT "Mandibuloacral dysplasia is caused by a mutation in LMNA-encoding  
 RT lamin A/C";  
 RL Am. J. Hum. Genet. 71:426-431(2002).  
 RN [16]  
 RP VARIANT CMD1A ASSOCIATED WITH QUADRIPEPS MYOPATHY HIS-377.  
 RX MEDLINE=22560388; PubMed=12673789; DOI=10.1002/humu.10170;  
 RA Charniot J.-C., Pascal C., Bouchier C., Sebillon P., Salama J.,  
 RA Duboscq-Bidot L., Peuchmaud M., Desnos M., Artigou J.-Y., Komajda M.,  
 RT "Functional consequences of an LMNA mutation associated with a new  
 RT cardiac and non-cardiac phenotype";  
 RL Hum. Mutat. 21:473-481(2003).  
 RN [17]  
 RP VARIANT LHCP LEU-133.  
 RX MEDLINE=22516017; PubMed=12625077; DOI=10.1210/jc.2002-021506;  
 RA Caux F., Duboscq-Bidot L., Lascos O., Buendia B., Chazouilleres O.,  
 RA Cohen A., Courvalin J.-C., Laroche L., Capeau J., Vigouroux C.,  
 RA Christin-Maitre S.;  
 RT "A new clinical condition linked to a novel mutation in lamins A and C  
 RT with generalized lipodystrophy, insulin-resistant diabetes, disseminated  
 RT leukodermic papules, liver steatosis and cardiomyopathy";  
 RL J. Clin. Endocrinol. Metab. 88:1006-1013(2003).  
 RN [18]  
 RP VARIANTS HGPS CYS-471; CYS-527 AND SER-608.  
 RX PubMed=12768443; DOI=10.1007/s10038-003-0025-3;  
 RA Cao H., Hegde R.A.;  
 RT "LMNA is mutated in Hutchinson-Gilford progeria (MIM 176670) but not  
 RT in Wiedemann-Rautenstrauch progeroid syndrome (MIM 264090)";  
 RL J. Hum. Genet. 48:271-274(2003).  
 RN [19]  
 RP VARIANT ATFB LYS-161.  
 RX PubMed=12920062;  
 RA Sebillon P., Bouchier C., Bidot L.D., Bonne G., Ahamed K., Charron P.,  
 RA Drouin-Garraud V., Millaire A., Desrumeaux G., Benache A.,  
 RA Charniot J.-C., Schwartz K., Villard E., Komajda M.;  
 RT "Expanding the phenotype of LMNA mutations in dilated cardiomyopathy  
 RT and functional consequences of these mutations";  
 RL J. Med. Genet. 40:560-567(2003).  
 RN [20]  
 RP VARIANTS WERNER SYNDROME PRO-57 AND ARG-140.  
 RX PubMed=12927431; DOI=10.1016/S0140-6736(03)14069-X;  
 RA Chen L., Lee L., Kudlow B.A., Dos Santos H.G., Sletvold O.,  
 RA Shafeghati Y., Botha E.G., Garg A., Hanson N.B., Martin G.M.,  
 RA Mian I.S., Kennedy B.K., Oshima J.;  
 RT "LMNA mutations in atypical Werner's syndrome";  
 RL Lancet 362:440-445(2003).  
 RN [21]  
 RP VARIANTS HGPS LYS-145 AND SER-608.  
 RX PubMed=12714972; DOI=10.1038/nature01629;  
 RA Eriksson M., Brown W.T., Gordon L.B., Glynn M.W., Singer J., Scott L.,  
 RA Erdos M.R., Robbins C.M., Moses T.Y., Berglund P., Dutra A., Pak E.,  
 RT Query Match 100.0%; Score 81; DB 1; Length 664;  
 RT Best Local Similarity 100.0%; Pred. No. 0.00014;  
 RT Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LLGNSPRTQSPNC 15  
 DB 647 LLGNSPRTQSPNC 661  
 RESULT 2  
 ID LAMA RAT  
 AC P48679; STANDARD; PRT; 665 AA.  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)

05-JUL-2004 (Rel. 44, Last annotation update)  
 Name=Lama; Synonyms=Lmn1;  
 Rattus norvegicus (Rat).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=93050186; PubMed=1426247; DOI=10.1016/0014-5793(92)80927-9;  
 Ozaki T., Sakiyama S.;  
 "Lamin A gene expression is specifically suppressed in v-src-  
 transformed cells";  
 RL FEBS Lett. 312:165-168(1992).  
 [2]  
 SEQUENCE OF 26-663 FROM N.A.  
 STRAIN=Sprague-Dawley; TISSUE=Liver;  
 Jommalagadda V.S., Farnak V.K.;  
 Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
 -1- FUNCTION: Lamins are components of the nuclear lamina, a fibrous  
 layer on the nucleoplasmic side of the inner nuclear membrane,  
 which is thought to provide a framework for the nuclear envelope  
 and may also interact with chromatin.  
 -1- SUBCELLULAR LOCATION: Nuclear.  
 -1- ALTERNATIVE PRODUCTS:  
 Event=Alternative splicing; Named isoforms=2;  
 Comment=Isoform A and isoform C are present in equal amounts in  
 the lamina of mammals;  
 Name=Lamin A;  
 IsoId=P48679-1; Sequence=Displayed;  
 Name=Lamin C;  
 IsoId=P48679-2; Sequence=Not described;  
 -1- PTM: Increased phosphorylation of the lamins occurs before  
 envelope disintegration and probably plays a role in regulating  
 lamin associations.  
 -1- MISCELLANEOUS: The structural integrity of the lamina is strictly  
 controlled by the cell cycle, as seen by the disintegration and  
 formation of the nuclear envelope in prophase and telophase,  
 respectively.  
 -1- SIMILARITY: Belongs to the intermediate filament family.  
 -1- CAUTION: Ref.1 sequence differs from that shown in the N- and C-  
 terminal due to frameshifts.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 EMBL; X66870; CAA47342.1; ALT\_FRAME.  
 EMBL; X76297; CAA53945.1; -.  
 HSSP; P02545; 11FR.  
 RGD; 620456; Lmna.  
 InterPro; IPR001664; IF.  
 InterPro; IPR001322; IF\_tail\_C.  
 Pfam; PF00038; Filament; 1.  
 Pfam; PF00932; IF\_tail; 1.  
 PROSITE; PS00226; IF; 1.  
 Alternative splicing; Coiled coil; Intermediate filament; Lipoprotein;  
 Nuclear protein; Phosphorylation; Prenylation.  
 DOMAIN 1 33 Head.  
 DOMAIN 34 383 Rod.  
 DOMAIN 384 665 Tail.  
 DOMAIN 34 70 Coil 1A.  
 DOMAIN 71 80 Linker 1.  
 DOMAIN 81 218 Coil 1B.  
 DOMAIN 219 242 Linker 2.  
 DOMAIN 243 383 Coil 2.  
 SITE 325 325 Stutter (By similarity).  
 DOMAIN 417 422 Nuclear localization signal (Potential).  
 LIPID 662 S-farnesyl cysteine (By similarity).

```

FT CONFLICT 470 470 X -> R (in Ref. 1).
FT CONFLICT 524 524 T -> S (in Ref. 1).
FT CONFLICT 584 584 R -> P (in Ref. 1).
FT CONFLICT 606 606 A -> P (in Ref. 1).
SQ SEQUENCE 665 AA; 74323 MW; 9CC553005C8534B4 CRC64;

Query Match 90.1%; Score 73; DB 1; Length 665;
Best Local Similarity 93.3%; Pred. No. 0.0029;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQNC 15
Db 648 LLGNSPRTQSQNC 662

RESULT 3
Q8BNX5 PRELIMINARY; PRT; 135 AA.
AC Q8BNX5;
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus adult male aorta and vein cDNA, RIKEN full-length
DE enriched library, clone: A530030G15 product: lamin A, full insert
DE sequence. (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20493974; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RIGA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
```

```

RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kurita H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK079972; BAC37795.1; -.
DR HSPB; P02545; IIVT. -.
FT NON TER 1
SQ SEQUENCE 135 AA; 13855 MW; AAES5854D9242B0 CRC64;

Query Match 85.2%; Score 69; DB 2; Length 135;
Best Local Similarity 86.7%; Pred. No. 0.0022;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQNC 15
Db 118 LLGNSPRTQSQNC 132

RESULT 4
LAMA_MOUSE STANDARD; PRT; 665 AA.
ID LAMA_MOUSE
AC P48678; P97859;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lamin A.
GN Name=Lma, Synonym=Lmnl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95300954; PubMed=7781761; DOI=10.1016/0014-5793(95)00453-G;
RA Nakajima N., Abe K.;
RT "Genomic structure of the mouse A-type lamin gene locus encoding
RT somatic and germ cell-specific lamins.";
RL FEBS Lett. 365:108-114(1995).
RN [2]
RP SEQUENCE OF 1-568 FROM N.A.
RX MEDLINE=8924740; PubMed=2719959; DOI=10.1016/0167-4781(89)90179-6;
RA Riedel W., Werner D.;
RT "Nucleotide sequence of the full-length mouse lamin C cDNA and its
RT deduced amino-acid sequence.";
RL Biochim. Biophys. Acta 1008:119-122(1989).
RN [3]
RP SEQUENCE OF 235-665 FROM N.A.
RX MEDLINE=93144345; PubMed=7916626; DOI=10.1016/0167-4781(93)90072-L;
RA Nakajima N., Sado T.;
RT "Nucleotide sequence of a mouse lamin A cDNA and its deduced amino
RT acid sequence.";
RL Biochim. Biophys. Acta 1171:311-314(1993).
CC -!- FUNCTION: Lamins are components of the nuclear lamina, a fibrous
CC layer on the nucleoplasmic side of the inner nuclear membrane,
CC which is thought to provide a framework for the nuclear envelope
CC and may also interact with chromatin.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Isoform A and isoform C are present in equal amounts in
CC the lamina of mammals;
```





```

RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Maehiro Y., Nagai K., Isegai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128329; BAC87386.1; -.
SQ SEQUENCE 172 AA; 19030 MW; 93988457BC63AC2E CRC64;

Query Match 59.3%; Score 48; DB 2; Length 172;
Best Local Similarity 76.9%; Pred. No. 7.8;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQ 13
Db 108 VLGYSPPRTLSFQ 120

RESULT 9
Q6CG31 PRELIMINARY; PRT; 852 AA.
ID Q6CG31
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similarity.
GN ORFNames=VALI0B01320g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA DuJon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neveuglise C., Talia E.,
RA Goffard N., Franchin L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boirame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.P., Straub M.L., Suleau A.,
RA Swennene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Winkler P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN SEQUENCE FROM N.A.
RP STRAIN=CLIB99;
RC Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382128; CAG82598.1; -.
DR InterPro; IPR008941; TPR-like.
SQ SEQUENCE 852 AA; 96983 MW; 178484411DE6302B CRC64;

Query Match 59.3%; Score 48; DB 2; Length 852;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LGNSPRTQSPQNC 15
Db 313 VGNLSPGQBDPQNC 326

RESULT 10

```

```

HUTH_PSEFL STANDARD; PRT; 514 AA.
ID HUTH_PSEFL
AC Q6WMR3;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine ammonia-lyase (EC 4.3.1.3) (Histidase).
GN Name=huth; Synonyms=huth-1;
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN SEQUENCE FROM N.A.
RC STRAIN=SBW25;
RA Zhang X.X.;
RT "Using RIVET to study histidine availability in planta for plant
RT growth-promoting Pseudomonas fluorescens SBW25";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: L-histidine = urocanate + NH(3).
CC -I- PATHWAY: Histidine degradation; first step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -I- PTM: Contains an active site 4-methylidene-imidazole-5-one (MIO),
CC which is formed autocatalytically by cyclization and dehydration
CC of residues Ala-Ser-Gly (By similarity).
CC -I- SIMILARITY: Belongs to the PAL / histidase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; AJ421809; CAD19072.1; -.
DR HSP; P21310; 1GKM.
DR HAMAP; MF 00229; -.
DR InterPro; IPR005921; Huth.
DR InterPro; IPR008948; L-Aspartase-like.
DR InterPro; IPR001106; Phe/His_NH3lyase.
DR Pfam; PF00221; PAL; 1.
DR TIGRFAMs; TIGR01225; huth; 1.
DR PROSITE; PS00488; PAL_HISTIDASE; 1.
DR Histidine metabolism; Lyase.
KW CROSSLINK 143 145 5-imidazolinone (Ala-Gly) (By
FT MOD_RES 144 144 similarity).
FT 2,3-didehydroalalanine (Ser) (By
FT similarity).
FT SEQUENCE 514 AA; 54478 MW; B89C067DA6985492 CRC64;

Query Match 56.8%; Score 46; DB 1; Length 514;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQNC 15
Db 260 LLGESSQVSHQNC 274

RESULT 11
Q9FWN9 PRELIMINARY; PRT; 162 AA.
ID Q9FWN9
AC Q9FWN9; Q7XCD4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative lipid transfer protein.
GN ORFNames=OSJNBa0015J15.17;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.

```

```

OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Uterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569(2003).
RN [4]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC026758; AAG3495.1; -.
DR EMBL; AE017118; AAP54939.1; -.
DR HSSP; P24337; 1HVP.
DR Gramene; Q7XCD4; -.
DR Gramene; Q9FWN9; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI1.1.
SQ SEQUENCE 162 AA; 16011 MW; 77A46824988B4086 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 162;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQNC 15
|||||:|
Db 93 LLGNSPSPSGEQCC 107

RESULT 12
Q8BPZ3 PRELIMINARY; PRT; 180 AA.
AC Q8BPZ3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mus musculus 12 days embryo eyeball cDNA, RIKEN full-length enriched
DE library, clone:D230011M17 product:hypothetical protein, full insert
DE sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Eyeball;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX Carninci P., Hayashizaki Y.;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Mech. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Eyeball;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Eyeball;
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX Oka T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,

```

```

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Eyeball;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Eyeball;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Eyeball;
RC STRAIN=C57BL/6J; TISSUE=Eyeball;
RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK051860; BAC34790.1; -.
KW Hypothetical protein.
SQ SEQUENCE 180 AA; 18945 MW; F839B5C2CE455504 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 180;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 PRQTSPQNC 15
|||||:|
Db 48 PRQTQAPQRC 56

RESULT 13
Q96MH2 PRELIMINARY; PRT; 286 AA.
ID Q96MH2
AC Q96MH2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein FLJ32384 (L3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Skeletal muscle;
RC PubMed=14702039; DOI=10.1038/ng1285;
RX Oka T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,

```

RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,  
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
RA Tanai H., Kinata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,  
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,  
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,  
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
RA Yoshishino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Mushioka Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,  
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shizu F., Wabe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,  
RT "Complete sequencing and characterization of 21,243 full-length human  
RT cDNAs.";  
RL Nat. Genet. 36:40-45 (2004).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Eye;  
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.J., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.A., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywicki M.I., Skalska U., Smalish U., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Eye;  
RA Strausberg R.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Petrozello J., Carter P.;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK056946; BAB71319.1; -;  
DR EMBL; BC025970; AAH25970.1; -;  
DR EMBL; AY598322; AAT06733.1; -;  
SQ SEQUENCE 286 AA; 32418 MW; 58576D72096A8A6F CRC64;

Query Match 54.3%; Score 44; DB 2; Length 286;  
Best Local Similarity 80.0%; Pred. No. 61;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 SPRTSQPNC 15  
|||||||

Db 71 SPRTSQPGC 80  
RESULT 14  
ID Q9VLOQ PRELIMINARY; PRT; 443 AA.  
AC Q9VLOQ;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE CG7851-PA (RH11377P).  
GN Name=Scgalpha; ORFNames=CG7851; fly).  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA DePamphilis M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
RA Fowler C., Gabrielian A.E., Garg N.F., Gelbart W.N., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.S., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laake B., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
RT melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
RT melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;



RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.

RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RG WormBase Consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; U80443; AAK68201.1; -;  
 DR HSSP; P11171; 1G33.  
 DR WormBase; WBGene0001491; fgm-4.  
 DR WormPep; C2A11.8b; CE27724.  
 DR GO; GO:0005856; C:cytoskeleton; IEA.  
 DR InterPro; IPR000299; Band 4.1.  
 DR InterPro; IPR009065; FERM.  
 DR InterPro; IPR011036; PH related.  
 DR Pfam; PF00373; Band 41; 1.  
 DR PRINTS; PR00935; BANO41.  
 DR SMART; SM00295; B41; 1.  
 DR PROSITE; PS0057; FERM\_3; 1.  
 SQ SEQUENCE 643 AA; 72668 MW; F7C1058404EA7182 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 643;  
 Best Local Similarity 61.5%; Pred. No. 1.5e+02;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSPRTQSPQ 13

Db 40 IIGNEFPRTQSAQ 52

RESULT 17

Q8H852 PRELIMINARY; PRT; 826 AA.  
 AC Q8H852;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Putative Nonclathrin coat protein gamma-like protein.  
 GN Name=OJ1626B05.5;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 ON NCBI TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,  
 RA Currie J., Collura K.;  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AC104473; AAN60990.1; -;  
 DR HSSP; Q9Y678; 1R4X.  
 DR Gramene; Q8H852; -;  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR InterPro; IPR002553; Adaptin\_N.  
 DR InterPro; IPR008938; ARM.  
 DR Pfam; PF01602; Adaptin\_N; 1.  
 SQ SEQUENCE 826 AA; 91594 MW; 2E44E7BB3BA8B61E CRC64;

Query Match 54.3%; Score 44; DB 2; Length 826;  
 Best Local Similarity 66.7%; Pred. No. 2e+02;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLGNSPRTQSP 12

Db 403 LLGNEGPRTSQP 414

RESULT 18

PSUI\_YEAST  
 ID PSUI\_YEAST STANDARD; PRT; 970 AA.  
 AC P53550;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE mRNA decapping protein 2 (PSUI protein).  
 GN Name=DCP2; Synonyms=PSUI; OrderedLocusNames=YNL118C; ORFNames=N1917;  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D273-10B;  
 RA Tzagoloff A.A.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97245296; PubMed=9090055;  
 RX DOI=10.1002/(SICI)1097-0061(19970315)13:3<261::AID-YEA64>3.0.CO;2-L;  
 RA de Antoni A., D'Angelo M., Dal Pero F., Sartorello F., Pandolfo D.,  
 RA Pallavicini A., Ianfranchi G., Valle G.;  
 RT "The DNA sequence of cosmid 14-13b from chromosome XIV of  
 RT Saccharomyces cerevisiae reveals an unusually high number of  
 RT overlapping open reading frames.";  
 RL Yeast 13:261-266(1997).  
 RN [3]  
 RP INTERACTION WITH DCP1.  
 RX MEDLINE=99438017; PubMed=10508173; DOI=10.1093/emboj/18.19.5411;  
 RA Duncley T., Parker R.;  
 RT "The DCP2 protein is required for mRNA decapping in Saccharomyces  
 RT cerevisiae and contains a functional MutT motif.";  
 RL EMBO J. 18:5411-5422(1999).  
 CC -!- FUNCTION: Required for the production of active decapping enzyme,  
 CC perhaps in a process requiring the hydrolysis of a pyrophosphate  
 CC bond. Decapping is the major pathway of mRNA degradation in yeast.  
 CC It occurs through deadenylation, decapping and subsequent 5' to 3',  
 CC exonucleolytic decay of the transcript body.  
 CC -!- SUBUNIT: Interacts with DCP1.  
 CC -!- SIMILARITY: Belongs to the Nudix hydrolase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L43065; AAA68866.1; -;  
 DR EMBL; Z69382; CAA93389.1; -;  
 DR EMBL; Z71394; CAA95998.1; -;  
 DR PIR; S63059; S63059.  
 DR GerMOnline; 143124; -;  
 DR SGD; S00005062; DCP2.  
 DR GO; GO:0000932; C:cytoplasmic mRNA processing body; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0008047; F:enzyme activator activity; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0003729; F:mRNA binding; IPI.  
 DR GO; GO:0000290; P:deadenylation-dependent decapping; IEA.  
 DR InterPro; IPR007722; DCP2.  
 DR InterPro; IPR000086; NUDIX\_hydrolase.  
 DR Pfam; PF05026; DCP2; 1.  
 DR Pfam; PF00293; NUDIX; 1.  
 DR PRINTS; PR00502; NUDIXFAMILY.  
 DR PROSITE; PS00893; NUDIX; 1.  
 DR Hydrolase.  
 KW DOMAIN 134 155 Nudix box.  
 FT DOMAIN 436 439 Poly-Ser.  
 FT CONFLICT 425 425 P -> L (in Ref. 1).  
 FT SEQUENCE 970 AA; 108667 MW; D53CA2CA546FA4A CRC64;

```

Query Match          54.3%; Score 44; DB 1; Length 970;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPON 14
Db 595 GKSSPSTQSKQN 606

RESULT 19
Q6C764 PRELIMINARY; PRT; 559 AA.
AC Q6C764;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Similar to sp|P38140 Saccharomyces cerevisiae YBR239c.
GN ORFNames=YAL10E03410g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RG STRAIN=CLIB99;
RA DuJon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisarane A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Kerrest F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swemene D., Tekaia F., Wesolowski-louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Boucher C., Caudron B., Scarpetti C., Gaillardin C., Weissenbach J.,
RA Winkler P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RG STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nucleus (By similarity).
CC -1- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
CC domain.
DR EMBL; CR382131; CAG79077.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0008270; P:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001138; Fungi_Trcsrp_N.
DR Pfam; PF00172; Zn_c1us; 1.
DR SMART; SM00066; GAL4; 1.
DR PROSITE; PS00048; ZN2_CV6_FUNGAL_2; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Transcription;
KW Transcription regulation; Zinc.
SQ SEQUENCE 559 AA; 61447 MW; DA57240C9E634D0B CRC64;

Query Match          53.1%; Score 43; DB 2; Length 559;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy 1 LLGNS--SPRTQSPON 14
Db 235 LLGNSQSPNTHSPHN 250

RESULT 21
Q9DSW7 PRELIMINARY; PRT; 161 AA.
AC Q9DSW7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Hypothetical protein.
OS Ascovirus DpAV4.
OC Viruses; dsDNA viruses, no RNA stage; Ascoviridae; Ascovirus.
OX NCBI_TaxID=113365;
RN [1]

```

## RESULT 20

```

Q8RZN9 PRELIMINARY; PRT; 2001 AA.
AC Q8RZN9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Putative polyprotein.
GN Name=B1065E10.26;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakaehima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yanagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
RL Gramene; Q8RZN9; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005162; Retrotrans_gag.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03732; Retrotrans_gag; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; RVT_1; 1.
KW Polyprotein; RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 2001 AA; 225569 MW; 93E5FE01BBFD7 CRC64;

Query Match          53.1%; Score 43; DB 2; Length 2001;
Best Local Similarity 57.1%; Pred. No. 7.7e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPON 14
Db 231 ILQSNIPRTQPPQN 244

RESULT 21
Q9DSW7 PRELIMINARY; PRT; 161 AA.
AC Q9DSW7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Hypothetical protein.
OS Ascovirus DpAV4.
OC Viruses; dsDNA viruses, no RNA stage; Ascoviridae; Ascovirus.
OX NCBI_TaxID=113365;
RN [1]

```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=20540044; PubMed=11086137;
RA Stacia K., Demattei M.V., Federici B.A., Bigot Y.;
RT "Phylogenetic position of the Diadromus pulchellus ascovirus DNA
RL polymerase among viruses with large double-stranded DNA genomes.";
RL J. Gen. Virol. 81:3059-3072(2000).
DR EMBL; AJ279812; CAC19118.1; -.
KW Hypothetical protein.
SQ SEQUENCE 161 AA; 18004 MW; DDBE914EFA8D2F4 CRC64;

Query Match 51.9%; Score 42; DB 2; Length 161;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GNSSPRTSQPN 14
Db 105 GASQETSPQN 116
| | | | |
| | | | |

RESULT 22
Q38966 PRELIMINARY; PRT; 191 AA.
AC Q38966;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=97086699; PubMed=8932388; DOI=10.1093/nar/24.21.4313;
RA Quigley F., Dao P., Cottet A., Mache R.;
RT "Sequence analysis of an 81 kb contig from Arabidopsis thaliana
RT chromosome III";
RL Nucleic Acids Res. 24:4313-4318(1996).
RN [2]
RN SEQUENCE FROM N.A.
RA Quigley F.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X98130; CAA66823.1; -.
KW Hypothetical protein.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20836 MW; A6CD3CAP835E1541 CRC64;

Query Match 51.9%; Score 42; DB 2; Length 191;
Best Local Similarity 61.5%; Pred. No. 83;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LGNSSPRTSQPN 14
Db 105 LGNSAPISPPN 117
| | | | |
| | | | |

RESULT 23
Q7D016 PRELIMINARY; PRT; 291 AA.
AC Q7D016;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR C 1564P.
OS OrderedLocNames=AGR C 1564;
OC Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Cereon;
```

```
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkley G., Gattung S., Miller N., Blanchard M.,
RA Quorllo B., Gorman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houmial K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Planagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AB080818; AAK86662.1; -.
DR InterPro; IPR004360; Gly_bico_diox.
DR Pfam; PF09093; Glyoxalase; 2.
SQ SEQUENCE 291 AA; 32554 MW; 233F81B8991FC496 CRC64;

Query Match 51.9%; Score 42; DB 2; Length 291;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 LGNSSPRTSQPN 13
Db 18 LSHTTPTQSP 29
| | | | |
| | | | |

RESULT 24
PERO DROME
ID _PERO DROME STANDARD; PRT; 690 AA.
AC Q01603; OSIJW8; Q7KSF9;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Peroxidase precursor (EC 1.11.1.7) (DmPO).
GN Name=Pxd; Synonym=HDC14047; PO; ORFNames=CG3477;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93129626; PubMed=1482687; DOI=10.1016/0167-4781(92)90127-L;
RA Ng S.W., Wiedemann M., Kontermann R., Petersen G.;
RT "Molecular characterization of a putative peroxidase gene of
RT Drosophila melanogaster.";
RL Biochim. Biophys. Acta 1171:224-228(1992).
RN [2]
RN SEQUENCE FROM N.A.
RA Konstandi O., Haeen Z., Stravopodis D., Papasideri I., Wever R.,
RA Margaritis L.H.;
RT "First functional expression of the chorion peroxidase in Pichia
RT pastoris";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherz S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
```

RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RA "The genome sequence of *Drosophila melanogaster*."  
RA Science 287:2185-2195(2000).  
RA [4]  
RN GENOME REANNOTATION.  
RP MEDLINE=22426069; PubMed=12537572;  
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.F.,  
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review."  
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [5]  
RN GENOME REANNOTATION.  
RP PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;  
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovyyev V., Busold C.,  
RA Fellenberg K., Boutsos M., Vingron M., Sauer F., Hoheisel J.D.,  
RA Paro R.;  
RT "An integrated gene annotation and transcriptional profiling approach  
RT towards the full gene content of the *Drosophila* genome."  
RL Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003).  
CC -!- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.  
CC -!- COPACTOR: Binds 1 heme B (iron-protoporphyrin IX) group per  
CC subunit [by similarity].  
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -!- SIMILARITY: Belongs to the peroxidase family. XPO subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X68131; CAA48238.1; -;  
DR EMBL; AY541497; AAS48542.1; -;  
DR EMBL; AE003715; AAS65161.1; -;  
DR EMBL; BK002598; DAA04104.1; ALT\_INIT.  
DR PIR; S28222; S28222.  
DR HSSP; P05164; ID5L.  
DR Flybase; FBgn0004577; Pxd.  
DR InterPro; IPR002007; Anim\_peroxidase.  
DR InterPro; IPR002016; Peroxidase.  
DR InterPro; IPR010255; Peroxidase\_super.  
DR Pfam; PF03098; An\_peroxidase\_1.  
DR PRINTS; PR00457; ANPEROXIDASE.  
DR PROSITE; PS00435; PEROXIDASE\_1; FALSE NEG.  
DR PROSITE; PS00436; PEROXIDASE\_2; FALSE NEG.  
DR PROSITE; PS50292; PEROXIDASE\_3; 1.  
KW Glycoprotein; Heme; Hydrogen peroxide; Iron; Oxidoreductase;

KW Peroxidase; Signal. 20  
FT SIGNAL 1 20 Potential.  
FT CHAIN 21 690 Peroxidase.  
FT ACT\_SITE 185 331 Proton acceptor (By similarity).  
FT ACT\_SITE 331 331 Charge stabilization (By similarity).  
FT METAL 437 437 Iron (heme axial ligand) (By similarity).  
FT METAL 437 437 N-linked (GlcNAc...) (Potential).  
FT COBOHYD 310 310 T -> A (in Ref. 3 and 5).  
FT CONFLICT 584 584 P -> L (in Ref. 3 and 5).  
FT CONFLICT 678 678  
SQ SEQUENCE 690 AA; 76733 MW; 4B16CE5411EB73C0 CRC64;  
Query Match 51.9%; Score 42; DB 1; Length 690;  
Best Local Similarity 50.0%; Pred. No. 3.4e+02;  
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LLGNSPRTSPQN 14  
:|||||:|  
Db 55 VIGNSLPTSPAPQN 68  
RESULT 25  
Q9LSE3 PRELIMINARY; PRT; 772 AA.  
ID Q9LSE3  
AC Q9LSE3;  
DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
DE Emb|CAA68822.1 (Putative RING zinc finger protein).  
GN Name=At3g26730;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OC NCBI\_TaxID=3702;  
OX [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20277480; PubMed=10819329;  
RA Nakamura Y.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC  
RT clones."  
RL DNA Res. 7:131-135(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;  
RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,  
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,  
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,  
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,  
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,  
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
DR EMBL; AB026648; BAB01736.1; -;  
DR EMBL; AY074325; AAL67021.1; -;  
DR EMBL; BT003004; AAO22812.1; -;  
DR HSSP; P38398; IJM7.  
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.

```

DR GO:0008270; P:zinc ion binding; IEA.
DR GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 772 AA; 85232 MW; B42710BEP868C45 CRC64;

Query Match 51.9%; Score 42; DB 2; Length 772;
Best Local Similarity 61.5%; Pred. No. 3.9e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LGNSSPRTQSPON 14
DB 667 LGNSAPISSSPPN 679

RESULT 26
QBY30
ID Q6BY30 PRELIMINARY; PRT; 901 AA.
AC Q6BY30;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Debaryomyces hansenii chromosome A of strain CBS767 of Debaryomyces
DE hansenii.
GN ORFNames=DEHA0A1321g;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Debaryomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
OX NCBI_TaxID=284592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aglie M., Anthouard V., Babour A., Barbé V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boismare A., Boyer J., Cattolico L., Confanioleti F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RA "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
CC domain.
DR EMBL; CR382133; CAG84866.1; -.
DR GO:0005634; C:nucleus; IEA.
DR GO:0003700; F:transcription factor activity; IEA.
DR GO:0008270; F:zinc ion binding; IEA.
DR GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO:0006350; P:transcription; IEA.
DR InterPro; IPR007219; Fungal_trans.
DR InterPro; IPR001138; Fungi_trscrp_N.
DR Pfam; PF04082; Fungal_trans; 1.
DR Pfam; PF00172; Zn_clus; 1.
DR PRINTS; PR00054; FUNGALZNCYS.
DR SMART; SM00066; GAL4; 1.

DR PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR PROSITE; PS50048; ZN2_CY6_FUNGAL_2; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Transcription;
KW Transcription regulation; Zinc.
SQ SEQUENCE 901 AA; 103042 MW; 80DE2AACBED76253 CRC64;

Query Match 51.9%; Score 42; DB 2; Length 901;
Best Local Similarity 66.7%; Pred. No. 4.6e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GNSSPRTQSPON 14
DB 16 GNNHPTIYPQN 27

RESULT 27
SUUR_DROER
ID SUUR_DROER STANDARD; PRT; 962 AA.
AC P59597;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Suppressor of Underreplication protein.
GN Name=SuUR;
OS Drosophila erecta (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7220;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Yurlova A., Belyakin S.N., Makunin I.V., Zhimulev I.F.;
RA "Cloning of Drosophila erecta SuUR gene";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Required for underreplication of DNA, which is found in
CC many late replicating euchromatic regions of salivary gland
CC polytene chromosomes. Controls chromatin organization in polytene
CC chromosomes (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear. Binds to polytene chromosomes from
CC salivary glands; localized at late-replicating intercalary
CC heterochromatin and pericentric heterochromatin. Colocalizes with
CC many Polycomb Group proteins binding sites on polytene chromosomes
CC (By similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; AJ539550; CAD62567.1; -.
DR FlyBase; FBgn0045001; Dere\SuUR.
DR InterPro; IPR000637; A+T_hook.
DR InterPro; IPR000330; SNF2_N.
DR Pfam; PF00176; SNF2_N; 1.
DR SMART; SM00384; AT_hook; 1.
KW Chromatin regulator; Developmental protein; Nuclear protein.
SQ SEQUENCE 962 AA; 107466 MW; C79A4E2BC52EF9D6 CRC64;

Query Match 51.9%; Score 42; DB 1; Length 962;
Best Local Similarity 57.1%; Pred. No. 5e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 LGNSSPRTQSPON 15
DB 367 LVNKSPTKSKKC 380

RESULT 28
Q23863

```

```

ID Q23863 PRELIMINARY; PRT; 2150 AA.
AC Q23863;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine kinase A.
GN Name=dhka;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX NCBI_TaxID=44689;
RN [1]
RC STRAIN=AX4;
RX MEDLINE=96324397; PubMed=8670894;
RA Wang N., Shaulsky G., Escalante R., Loomis W.F.;
RT "A two-component histidine kinase gene that functions in Dictyostelium
development.";
RL EMOB J. 15:3890-3898(1996).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; U42597; AAC47300.1; -.
DR PIR; S71629; S71629.
DR HSSP; P39928; 10XK.
DR DictyBase; DB0215354; dhka.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0001630; F:kinase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR003594; AtPbind_Arpase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR006189; CHASE.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF03924; CHASE; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00512; Hiska; 1.
DR Pfam; PF00072; Response_reg; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS0839; CHASE; 1.
DR PROSITE; PS0109; HIS_KIN; 1.
DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction.
SQ SEQUENCE 2150 AA; 239662 MW; 1FC3F63CE0336FA2 CRC64;

Query Match 51.9%; Score 42; DB 2; Length 2150;
Best Local Similarity 53.8%; Pred. No. 1.2e+03;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 GNSSPRTQSPNC 15
DB 345 GNSSPRNCNGSNC 357

RESULT 29
OY Q7TPH6 PRELIMINARY; PRT; 4708 AA.
AC Q7TPH6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Highwire.
GN Name=Phrl;
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX NCBI_TaxID=10090;
RN [1]
RC STRAIN=C57BL/6J;
RX PubMed=1472956; DOI=10.1128/MCB.24.3.1096-1105.2004;
RA Burgess R.W., Peterson K.A., Johnson M.J., Roix J.J., Welsh I.C.,
O'Brien T.P.;
RT "Evidence for a conserved function in synapse formation reveals Phrl
as a candidate gene for respiratory failure in newborn mice.";
RL Mol. Cell. Biol. 24:1096-1105(2004).
DR EMBL; AY325887; AAP8591.1; -.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001298; Filamin.
DR InterPro; IPR009091; RCC1/BLIP-II.
DR InterPro; IPR000408; Reg_chrom_condens.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00415; RCC1; 1.
DR SMART; SM0184; RING; 1.
DR PROSITE; PS00150; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00194; FILAMIN_REPEAT; 1.
DR PROSITE; PS00626; RCC1_2; 1.
DR PROSITE; PS50012; RCC1_3; 3.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 4708 AA; 517357 MW; B9E8717F81DFD5F8 CRC64;

Query Match 51.9%; Score 42; DB 2; Length 4708;
Best Local Similarity 66.7%; Pred. No. 2.9e+03;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 GNSSPRTQSPQN 14
DB 2870 GASSPRSSSPQD 2881

RESULT 30
OY Q8YL47 PRELIMINARY; PRT; 148 AA.
AC Q8YL47;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alr7088 protein.
GN OrderedLocusNames=alr7088;
OS Anabaena sp. (strain PCC 7120).
OG Bacteria; Cyanobacteria.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RC STRAIN=PCC 7120;
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003600; BAB78172.1; -.
DR PIR; AH2488; AH2488.
KW Complete proteome; Plasmid.
SQ SEQUENCE 148 AA; 16201 MW; 895E89E2A0637911 CRC64;

Query Match 50.6%; Score 41; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```





OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Amorphaeae; Amorpha.  
 OK NCBI\_TaxID=48131;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Stem;  
 RA Liu W.-Z., An L.-J.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme  
 CC family.  
 DR EMBL; AF435968; AAL35216.1; -.  
 DR HSSP; P08659; 1LCI.  
 DR GO; GO:0016207; F:4-coumarate-CoA ligase activity; IEA.  
 DR GO; GO:0016874; P:ligase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR PRINTS; PR00154; AMPBINDING.  
 DR PROSITE; PS00455; AMP\_BINDING; 1.  
 KW Ligase.  
 SQ SEQUENCE 540 AA; 59315 MW; F399023E92440B79 CRC64;  
  
 Query Match 50.6%; Score 41; DB 2; Length 540;  
 Best Local Similarity 61.5%; Pred. No. 3.8e+02;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 3 GNSSPRTQSPQNC 15  
 ||||| :  
 DB 374 GNSLPRNQSGBIC 386  
  
 RESULT 38  
 4CL2\_TOBAC STANDARD; PRT; 542 AA.  
 ID 4CL2\_TOBAC  
 AC O24146;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE 4-coumarate-CoA ligase 2 (EC 6.2.1.12) (4CL 2) (4-coumaroyl-CoA  
 synthase 2).  
 GN Name=4CL2;  
 OS Nicotiana tabacum (Common tobacco).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 CC lamiales; Solanales; Solanaceae; Nicotiana.  
 OK NCBI\_TaxID=4097;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=96416441; PubMed=8819324; DOI=10.1104/pp.112.1.193;  
 RA Lee D., Douglas C.J.;  
 RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase  
 (4CL) gene family. cDNA structure, gene inheritance and expression,  
 and properties of recombinant proteins.";  
 RL Plant Physiol. 112:193-205(1996).  
 CC -!- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate +  
 CC 4-coumaroyl-CoA.  
 CC -!- PATHWAY: Branch-point reactions between general phenylpropanoid  
 CC metabolism and pathways leading to various specific end products.  
 CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U50846; AAB18638.1; -.  
 DR PIR; T03789; T03789.  
 DR HSSP; P08659; 1LCI.  
 DR InterPro; IPR000873; AMP-bind.

DR Pfam; PF00501; AMP-binding; 1.  
 DR PRINTS; PR00154; AMPBINDING.  
 DR TIGRFAMS; TIGR01733; AA-adenyl-dom; 1.  
 DR TIGRFAMS; TIGR01923; menE; 1.  
 DR PROSITE; PS00455; AMP\_BINDING; 1.  
 KW Ligase; Multigene family; Phenylpropanoid metabolism.  
 SQ SEQUENCE 542 AA; 59479 MW; CB5579AEDFCFC003 CRC64;  
  
 Query Match 50.6%; Score 41; DB 1; Length 542;  
 Best Local Similarity 61.5%; Pred. No. 3.8e+02;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 3 GNSSPRTQSPQNC 15  
 ||||| :  
 DB 377 GNSLPRNQSGBIC 389  
  
 RESULT 39  
 Q42943 PRELIMINARY; PRT; 542 AA.  
 ID Q42943  
 AC Q42943;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE 4-coumarate:coenzyme A ligase (EC 6.2.1.12).  
 OS Nicotiana tabacum (Common tobacco).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 CC lamiales; Solanales; Solanaceae; Nicotiana.  
 OK NCBI\_TaxID=4097;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=SRI;  
 RA Katayama Y., Kawai S., Morohoshi N., Kajita S.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme  
 CC family.  
 DR EMBL; D43773; BAA07828.1; -.  
 DR PIR; T02074; T02074.  
 DR HSSP; P08659; 1LCI.  
 DR GO; GO:0016207; F:4-coumarate-CoA ligase activity; IEA.  
 DR GO; GO:0016874; P:ligase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR000873; AMP-bind.  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR PRINTS; PR00154; AMPBINDING.  
 DR PROSITE; PS00455; AMP\_BINDING; 1.  
 KW Ligase.  
 SQ SEQUENCE 542 AA; 59450 MW; C4EEFACCC4EA0650 CRC64;  
  
 Query Match 50.6%; Score 41; DB 2; Length 542;  
 Best Local Similarity 61.5%; Pred. No. 3.8e+02;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 3 GNSSPRTQSPQNC 15  
 ||||| :  
 DB 377 GNSLPRNQSGBIC 389  
  
 RESULT 40  
 Q6TUF8 PRELIMINARY; PRT; 556 AA.  
 ID Q6TUF8  
 AC Q6TUF8;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE LRRGT00086.  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OK NCBI\_TaxID=10116;  
 RN [1]  
 RN SEQUENCE FROM N.A.

```

RC STRAIN=Sprague-Dawley;
RA Xu C.S., Chang C.F., Han H.P., Wang G.P., Chai L.O., Yuan J.Y.,
RA Yang K.J., Zhao L.F., Ma H., Wang L., Wang S.F., Xing X.K., Shen G.M.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY387072; AAQ91042.1; -.
DR HSSP; P00750; 1ASH.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001177; Apple.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00024; PAN; 3.
DR PRINTS; PR00005; APPLEDOMAIN.
DR SMART; SM00223; APPLE; 3.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00495; APPLE; 1.
DR PROSITE; PS50948; PAN; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 556 AA; 62989 MW; 0A15BCA5580E2558 CRC64;

Query Match 50.6%; Score 41; DB 2; Length 556;
Best Local Similarity 58.3%; Pred. No. 4e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 LGSSSPRTQSPQ I3
||:|||||:|:|
DB 392 LGSSSPSTETPK 403

RESULT 41
Q8PIR2 PRELIMINARY; PRT; 577 AA.
AC Q8PIR2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular serine protease.
GN OrderedLocustNames=XAC2833;
OC Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Cimarotti G., Cannavaro F., Cardoso J., Chambergo F., Cispina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Trinidade L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing host specificities."
RL Nature 417:459-463(2002).
CC -!- SIMILARITY: Belongs to peptidase family S8.

```

```

DR EMBL; AE011924; AAM37678.1; -.
DR HSSP; P00780; ICSE.
DR GO; GO:0004269; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008979; Gal bind like.
DR InterPro; IPR000209; Pept_S8_S53.
DR InterPro; IPR002884; Prptotnconvertap.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF01483; P_protein; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR ProDom; PD000717; Prptotnconvertsp; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; UNKNOWN 1.
KW Complete proteome; Hydrolase; Protease; Serine protease.
SQ SEQUENCE 577 AA; 61529 MW; 33777969B1BC37E CRC64;

Query Match 50.6%; Score 41; DB 2; Length 577;
Best Local Similarity 58.3%; Pred. No. 4.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 GNSSPRTQSPQN 14
||:|||||:|:|
DB 330 GNNRDRTRAPQN 341

RESULT 42
CDYL MOUSE
ID --CDYL MOUSE STANDARD; PRT; 593 AA.
AC Q9WTK2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chromodomain Y-like protein (CDY-like).
GN Name=Cdy1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=99206615; PubMed=10192397; DOI=10.1038/7771;
RA Lahn B.T., Page D.C.;
RT "Retroposition of autosomal mRNA yielded testis-specific gene family on human Y chromosome."
RL Nat. Genet. 21:429-433(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed as 2 transcripts: a ubiquitous

```



```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.I., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinchwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Director MGC Project;
RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Contains 1 SH3 domain.
DR EMBL; BC078655; AAT78655.1; -.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011036; PH related.
DR InterPro; IPR000219; RhogEF.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR011511; SH3_2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhogEF; 1.
DR Pfam; PF00018; SH3_1; 1.
DR Pfam; PF07653; SH3_2; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhogEF; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50010; PH DOMAIN; 1.
DR PROSITE; PS50002; SH3; 1.
DR SH3 domain.
KW SH3 domain.
SQ SEQUENCE 871 AA; 97347 MW; 43BB87127A05144B CRC64;

Query Match 50.6%; Score 41; DB 2; Length 871;
Best Local Similarity 64.3%; Pred. No. 6.5e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQN 14
Db 29 LLGSKPRPQSQS 42

RESULT 46
Q6Q8Q8 ID Q6Q8Q8 PRELIMINARY; PRT; 871 AA.
AC Q6Q8Q8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SH3-containing guanine nucleotide exchange factor.
GN Name=SGEF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=15133129;
RA Ellerbroek S.M., Wennerberg K., Arthur W.T., Dunty J.M., Bowman D.R.,
RA DeMali K.A., Der C., Burridge K.;
RT "SGEF, a Rhog Guanine Nucleotide Exchange Factor that Stimulates
RT Macropinocytosis.";
RL Mol. Biol. Cell 15:3309-3319(2004).
CC -I- SIMILARITY: Contains 1 PH domain.
CC -I- SIMILARITY: Contains 1 SH3 domain.
DR EMBL; AY552599; AAS59842.1; -.
DR HSP; P19878; IK4U
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhogEF.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhogEF; 1.
DR Pfam; PF00018; SH3_1; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhogEF; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50010; PH DOMAIN; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50002; SH3; 1.
DR SH3 domain.
KW SH3 domain.
SQ SEQUENCE 871 AA; 97333 MW; 28D081A4A3898B61 CRC64;

Query Match 50.6%; Score 41; DB 2; Length 871;
Best Local Similarity 64.3%; Pred. No. 6.5e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQN 14
Db 29 LLGSKPRPQSQS 42

RESULT 47
Q6DRCS ID Q6DRCS PRELIMINARY; PRT; 1014 AA.
AC Q6DRCS;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE D3S106E-like.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15256591; DOI=10.1073/pnas.0403929101;
RA Amsterdam A., Nissen R.M., Sun Z., Swindell E.C., Farrington S.,
RA Hopkins N.;
RT "Identification of 315 genes essential for early zebrafish
RT development.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12792-12797(2004).
DR EMBL; AY648834; AAT68152.1; -.
SQ SEQUENCE 1014 AA; 112531 MW; 14BCCCA22B2BC5A6 CRC64;

Query Match 50.6%; Score 41; DB 2; Length 1014;
Best Local Similarity 57.1%; Pred. No. 7.7e+02;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 LGNSSPRTQSPQNC 15
Db 19 LGKSKDRNASPGNC 32

RESULT 48
FSH_DROME
```

ID PSH DROME STANDARD; PRT; 2038 AA.

AC P13709; P13710;

DT 01-JAN-1990 (rel. 13, Created)

DT 01-JAN-1990 (rel. 13, Last sequence update)

DT 05-JUL-2004 (rel. 44, Last annotation update)

DE Female sterile homeotic protein (Fragile-chorion membrane protein).

GN Name=fs(1)h; Synonyms=fs(1)h;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89276730; PubMed=2567251;

RA Haynes S.R., Mozer B.A., Bhatia-Dey N., Dawid I.B.;

RT "The Drosophila fish locus, a maternal effect homeotic gene, encodes

RT apparent membrane proteins.";

RL Dev. Biol. 134:246-257(1989).

CC -!- FUNCTION: Required maternally for proper expression of other

CC homeotic genes involved in pattern formation, such as Ubx.

CC -!- SIMILARITY: Contains 2 bromodomains.

CC -!- SIMILARITY: Contains 1 ET domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M23221; AAA28540.1; -

CC EMBL; M23222; AAA28541.1; ALT TERM.

CC EMBL; M15762; AAA70424.1; -

CC EMBL; M15763; AAA70423.1; -

CC EMBL; M15764; AAA70422.1; -

CC FIC; A43742; A43742.

CC HSP; Q92830; 1F68.

CC FlyBase; FBgn004656; fs(1)h.

CC InterPro; IPR001487; Bromodomain.

CC Pfam; PF00439; Bromodomain; 2.

CC PRINTS; PR00503; BROMODOMAIN.

CC SMART; SM00297; BROMO; 2.

CC PROSITE; PS00633; BROMODOMAIN\_1; 2.

CC PROSITE; PS50014; BROMODOMAIN\_2; 2.

CC Bromodomain; Developmental protein; Repeat; Transmembrane.

FT DOMAIN 51 123

FT DOMAIN 495 567

FT DOMAIN 945 1106

FT TRANSMEM 330 350

FT TRANSMEM 451 471

FT TRANSMEM 750 770

FT TRANSMEM 790 810

FT TRANSMEM 816 830

FT TRANSMEM 874 894

FT TRANSMEM 1731 1751

FT TRANSMEM 1939 1959

FT VARIANT 909 909

FT VARIANT 1022 1022

FT VARIANT 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

FT SEQUENCE

Query Match 50.6%; Score 41; DB 1; Length 2038;

Best Local Similarity 66.7%; Pred. NO. 1.7e+03;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GNSSPRTQSPQN 14

DB 1985 GOSPPAQSPQD 1996

RESULT 49

Q9W3L3

ID AC

DT DT

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE CG2252-PB.

GN Name=fs(1)h; ORFNames=CG2252;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brannon R.C., Rogers J.H., Blazell R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,

RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,

RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K.J., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,

RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,

RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Ye J.,

RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,

RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

RP SEQUENCE FROM N.A.

RX MEDLINE=22426065; PubMed=12537568;

RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,

RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,

RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,

RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;

RT "Finishing a whole-genome shotgun. Release 3 of the Drosophila

RT melanogaster euchromatic genome sequence.";

RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

RP SEQUENCE FROM N.A.

RX MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskaas R.,

RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,

RA Ashburner M., Celniker S.E.;

RT "The transposable elements of the Drosophila melanogaster euchromatin:

RT a genomics perspective.";

```
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=23426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berwan B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003442; AAF46312.3; -.
DR HSP; Q92830; 1F68.
DR IntAct; Q9W3L3; -.
DR FlyBase; FBgn004656; fs(1)h.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR002173; PfkB.
DR Pfam; PF00439; Bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN 1; 2.
DR PROSITE; PS50014; BROMODOMAIN 2; 2.
DR PROSITE; PS00583; PFKB KINASE 1; UNKNOWN 1.
SQ SEQUENCE 2038 AA; 205345 MW; DC4ALA7B1266191E CRC64;

Query Match 50.6%; Score 41; DB 2; Length 2038;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GNSSPRTSQPN 14
Db 1985 GQSSPAQQSPQD 1996

RESULT 50
SX12 XENLA STANDARD; PRT; 470 AA.
AC P40649;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE SOX-12 protein.
GN Name=SOX-12;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=96180312; PubMed=8597594; DOI=10.1016/0167-4781(95)00228-6;
RA Komatsu N., Hiraoka Y., Shiozawa M., Ogawa M., Aiso S.;
RT "Cloning and expression of Xenopus laevis xSox12 cDNA.";
RL Biochim. Biophys. Acta 1305:117-119(1996).
RN [2]
RP SEQUENCE OF 306-359 FROM N.A.
RX MEDLINE=92310993; PubMed=1614875;
RA Denny P., Swift S., Brand N., Dabhadre N., Barton P., Ashworth A.;
RT "A conserved family of genes related to the testis determining gene,
RT SYX.";
RL Nucleic Acids Res. 20:2887-2887(1992).
```

```
CC -i- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -i- SIMILARITY: Contains 1 HMG box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50552; BAA09119.1; -.
DR EMBL; X65655; CAA46606.1; -.
DR PIR; S71466; S71466.
DR HSP; P35710; I111.
DR InterPro; IPR009071; HMG-box.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
DR PROSITE; PS50118; HMG_BOX_2; 1.
KW DNA-binding; Nuclear protein. HMG box.
FT DNA_BIND 295 363
SQ SEQUENCE 470 AA; 52859 MW; D0FE3BE5A08406AE CRC64;

Query Match 50.0%; Score 40.5; DB 1; Length 470;
Best Local Similarity 66.7%; Pred. No. 4e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Qy 1 LLGNSSPRTSQPN 15
Db 415 LLGNLS---QSPQEC 426

Search completed: August 25, 2005, 11:43:52
Job time : 174 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 11:35:23 ; Search time 41 Seconds  
(without alignments)  
27.311 Million cell updates/sec

Title: US-10-803-541-2  
Perfect score: 81  
Sequence: 1 LLGNSSPRTQSPQNC 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	515	2	US-08-705-660-46
2	81	100.0	515	3	US-08-989-045-46
3	81	100.0	515	4	US-09-315-355A-46
4	81	100.0	664	4	US-09-538-092-838
5	44	54.3	83	4	US-09-621-976-4157
6	44	54.3	970	4	US-09-538-092-664
7	42	51.9	88	4	US-09-621-976-3949
8	42	51.9	88	4	US-09-621-976-5346
9	42	51.9	164	4	US-09-252-991A-30496
10	42	51.9	177	5	PCT-US96-03916-19
11	41	50.6	10	2	US-08-429-964-16
12	41	50.6	10	5	PCT-US93-08062-16
13	40	49.4	136	4	US-09-270-767-45686
14	40	49.4	1031	4	US-09-949-016-7849
15	39	48.1	144	4	US-09-252-991A-25786
16	39	48.1	281	1	US-07-800-364B-14
17	39	48.1	281	2	US-07-989-847-12
18	39	48.1	281	3	US-08-469-411-12
19	39	48.1	281	4	US-09-780-601A-12
20	39	48.1	399	5	PCT-US91-07635-6
21	39	48.1	402	1	US-07-841-646-29
22	39	48.1	402	1	US-07-901-703-11
23	39	48.1	402	1	US-08-147-023-29
24	39	48.1	402	1	US-08-206-864-4
25	39	48.1	402	1	US-08-278-729A-21
26	39	48.1	402	1	US-08-480-528A-8
27	39	48.1	402	1	US-08-479-666-8

39	48.1	402	1	US-08-155-343A-21	Sequence 21, Appl
39	48.1	402	1	US-08-406-672-21	Sequence 21, Appl
39	48.1	402	1	US-08-643-563A-21	Sequence 21, Appl
39	48.1	402	1	US-08-447-570-29	Sequence 29, Appl
39	48.1	402	1	US-08-643-763A-21	Sequence 21, Appl
39	48.1	402	1	US-08-462-623-21	Sequence 21, Appl
39	48.1	402	1	US-08-451-953A-21	Sequence 21, Appl
39	48.1	402	2	US-08-459-346-6	Sequence 6, Appl
39	48.1	402	2	US-08-445-468A-21	Sequence 21, Appl
39	48.1	402	2	US-08-901-200A-8	Sequence 8, Appl
39	48.1	402	2	US-08-449-700-29	Sequence 29, Appl
39	48.1	402	2	US-08-449-699A-29	Sequence 29, Appl
39	48.1	402	2	US-08-461-397A-21	Sequence 21, Appl
39	48.1	402	2	US-08-912-088-21	Sequence 21, Appl
39	48.1	402	2	US-08-278-730A-21	Sequence 21, Appl
39	48.1	402	3	US-08-889-419-6	Sequence 6, Appl
39	48.1	402	3	US-08-445-467-21	Sequence 21, Appl
39	48.1	402	3	US-08-480-515A-21	Sequence 21, Appl
39	48.1	402	3	US-09-219-391-8	Sequence 8, Appl
39	48.1	402	3	US-09-170-936-21	Sequence 21, Appl
39	48.1	402	3	US-08-402-542-6	Sequence 6, Appl
39	48.1	402	3	US-08-461-113-21	Sequence 21, Appl
39	48.1	402	4	US-08-456-033-21	Sequence 21, Appl
39	48.1	402	4	US-08-643-321-20	Sequence 20, Appl
39	48.1	402	4	US-09-148-925C-29	Sequence 29, Appl
39	48.1	402	4	US-08-957-425-29	Sequence 29, Appl
39	48.1	402	4	US-08-260-675-21	Sequence 21, Appl
39	48.1	402	5	PCT-US92-01968-21	Sequence 21, Appl
39	48.1	402	5	PCT-US93-05446-11	Sequence 11, Appl
39	48.1	402	5	PCT-US93-07189-6	Sequence 6, Appl
39	48.1	402	5	PCT-US93-07190-21	Sequence 21, Appl
39	48.1	402	5	PCT-US93-07231-21	Sequence 21, Appl
39	48.1	402	5	PCT-US93-08742-21	Sequence 21, Appl
39	48.1	402	5	PCT-US93-08808-21	Sequence 21, Appl
39	48.1	402	5	PCT-US93-08885-21	Sequence 21, Appl
39	48.1	402	5	PCT-US93-10520-8	Sequence 8, Appl
39	48.1	445	1	US-08-447-925-6	Sequence 6, Appl
39	48.1	452	4	US-09-949-016-9081	Sequence 9081, Ap
39	48.1	492	4	US-09-096-731A-2	Sequence 2, Appl
39	48.1	497	4	US-09-499-846-6	Sequence 6, Appl
39	48.1	497	4	US-09-499-846-10	Sequence 10, Appl
39	48.1	525	4	US-09-499-846-4	Sequence 4, Appl
39	48.1	525	4	US-09-499-846-8	Sequence 8, Appl
39	48.1	56	4	US-09-621-976-5720	Sequence 5720, Ap
39	48.1	66	4	US-09-513-999C-4533	Sequence 4533, Ap
39	48.1	66	4	US-09-471-276-965	Sequence 965, App
39	48.1	107	4	US-09-252-991A-22561	Sequence 22561, A
39	48.1	120	4	US-09-270-767-39913	Sequence 39913, A
39	48.1	120	4	US-09-270-767-55130	Sequence 55130, A
39	48.1	128	4	US-09-513-999C-5518	Sequence 5518, Ap
39	48.1	135	4	US-09-252-991A-24945	Sequence 24945, A
39	48.1	137	4	US-09-252-991A-17872	Sequence 17872, A
39	48.1	293	4	US-09-252-991A-29843	Sequence 29843, A
39	48.1	293	4	US-09-252-991A-27425	Sequence 27425, A
39	48.1	305	4	US-09-635-872A-3	Sequence 3, Appl
39	48.1	305	4	US-09-636-077A-3	Sequence 3, Appl
39	48.1	305	4	US-09-636-060C-3	Sequence 3, Appl
39	48.1	305	4	US-09-986-552C-3	Sequence 3, Appl
39	48.1	305	4	US-09-636-596C-3	Sequence 3, Appl
39	48.1	305	4	US-10-023-894-7	Sequence 7, Appl
39	48.1	305	4	US-10-306-686-3	Sequence 3, Appl
39	48.1	453	2	US-08-484-126-3	Sequence 3, Appl
39	48.1	453	2	US-09-374-909-3	Sequence 3, Appl
39	48.1	469	2	US-08-484-126-1	Sequence 1, Appl
39	48.1	469	4	US-09-374-909-1	Sequence 1, Appl
39	48.1	469	4	US-09-252-991A-30596	Sequence 30596, A
39	48.1	469	4	US-09-248-796A-22382	Sequence 22382, A
39	48.1	622	4	US-09-499-846-2	Sequence 2, Appl
39	48.1	632	4	US-09-315-127-2	Sequence 2, Appl
39	48.1	632	4	US-09-315-127-3	Sequence 3, Appl
39	48.1	665	4	US-09-309-572-14	Sequence 14, Appl
39	48.1	665	4	US-09-718-096-14	Sequence 14, Appl
39	48.1	864	4	US-09-883-096-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT 1
US-08-705-660-46
; Sequence 46, Application US/08705660
; Patent No. 5858683
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,660
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-705-660-46

Query Match 100.0%; Score 81; DB 2; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15
Db 498 LLGNSSPRTQSPQNC 512

RESULT 2
US-08-989-045-46
; Sequence 46, Application US/08989045
; Patent No. 6027905
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,045
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-989-045-46

Query Match 100.0%; Score 81; DB 3; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15
Db 498 LLGNSSPRTQSPQNC 512

RESULT 3
US-09-315-355A-46
; Sequence 46, Application US/09315355A
; Patent No. 6803189
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: Methods for the Detection of Cervical Cancer
; FILE REFERENCE: MTP-023DV2
; CURRENT APPLICATION NUMBER: US/09/315,355A
; CURRENT FILING DATE: 1999-05-17
; PRIOR APPLICATION NUMBER: US 08/989,045
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: US 08/705,660
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-355A-46

Query Match 100.0%; Score 81; DB 4; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15
Db 498 LLGNSSPRTQSPQNC 512

RESULT 4
US-09-538-092-838
; Sequence 838, Application US/09538092
; Patent No. 6753314

```

```

; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,045
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-989-045-46

Query Match 100.0%; Score 81; DB 3; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15
Db 498 LLGNSSPRTQSPQNC 512

RESULT 3
US-09-315-355A-46
; Sequence 46, Application US/09315355A
; Patent No. 6803189
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: Methods for the Detection of Cervical Cancer
; FILE REFERENCE: MTP-023DV2
; CURRENT APPLICATION NUMBER: US/09/315,355A
; CURRENT FILING DATE: 1999-05-17
; PRIOR APPLICATION NUMBER: US 08/989,045
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: US 08/705,660
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-355A-46

Query Match 100.0%; Score 81; DB 4; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15
Db 498 LLGNSSPRTQSPQNC 512

RESULT 4
US-09-538-092-838
; Sequence 838, Application US/09538092
; Patent No. 6753314

```

```
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 838
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P02545
;
US-09-538-092-838

Query Match      100.0%; Score 81; DB 4; Length 664;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQNC 15
Db 647 LLGNSSPRTQSPQNC 661

RESULT 5
US-09-621-976-4157
; Sequence 4157, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4157
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23...-1
US-09-621-976-4157

Query Match      54.3%; Score 44; DB 4; Length 83;
Best Local Similarity 64.3%; Pred. No. 5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPON 14
Db 13 LLGSSSPPTSASQN 26

RESULT 6
US-09-538-092-664
; Sequence 664, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
```

```
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 664
; LENGTH: 970
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number YNL118C
;
US-09-538-092-664

Query Match      54.3%; Score 44; DB 4; Length 970;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GNSSPRTQSPON 14
Db 595 GKSSPSTQSQNW 606

RESULT 7
US-09-621-976-3949
; Sequence 3949, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3949
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -37...-1
; NAME/KEY: UNSURE
; LOCATION: 10
; OTHER INFORMATION: Xaa = Arg,Thr
US-09-621-976-3949

Query Match      51.9%; Score 42; DB 4; Length 88;
Best Local Similarity 64.3%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPON 14
Db 27 LLGSSSPPTLASQN 40

RESULT 8
US-09-621-976-5346
; Sequence 5346, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
```

```

; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5346
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: -37...-1
; NAME/KEY: UNSURE
; LOCATION: 10
; OTHER INFORMATION: Xaa = Arg,Thr
US-09-621-976-5346

Query Match          51.9%; Score 42; DB 4; Length 89;
Best Local Similarity 64.3%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGSSPRTQSPQN 14
Db 27 LLGSSPPTLASQN 40

RESULT 9
US-09-252-991A-30496
; Sequence 30496, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30496
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30496

Query Match          51.9%; Score 42; DB 4; Length 164;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SSPRTQSPQC 15
Db 22 SNPRHQEQSC 32

RESULT 10
PCT-US96-03916-19
; Sequence 19, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-03916-19

Query Match          51.9%; Score 42; DB 5; Length 177;
Best Local Similarity 53.3%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LLGSSPRTQSPQC 15
Db 5 LLNRGSPRLNPPKC 19

RESULT 11
US-08-429-964-16
; Sequence 16, Application US/08429964
; Patent No. 5962243
; GENERAL INFORMATION:
; APPLICANT: BROWN, MICHAEL S.
; APPLICANT: GOLDSTEIN, JOSEPH L.
; APPLICANT: REISS, YUVAL
; APPLICANT: JAMES, GUY L.
; TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P. O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/429,964
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,625
; FILING DATE: 16-FEB-1993
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/822,011
; FILING DATE: ABANDONED
; CLASSIFICATION: 435
; APPLICATION NUMBER: PCT/US/91/02650
; FILING DATE: 18-APR-1991
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/615,715

```

```

; FILING DATE: 20-NOV-1990
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/510,706
; FILING DATE: 18-APR-1990 (ABANDONED)
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:432/PAR
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 785-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-429-964-16

Query Match 50.6%; Score 41; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TQSPQC 15
Db 1 TQSPQC 7

RESULT 12
PCT-US93-08062-16
; Sequence 16, Application PC/TUS9308062
; GENERAL INFORMATION:
; APPLICANT:
; SEQUENCE CHARACTERISTICS: BROWN, MICHAEL S.
; SEQUENCE CHARACTERISTICS: GOLDSTEIN, JOSEPH L.
; SEQUENCE CHARACTERISTICS: REIS, YUVAL
; SEQUENCE CHARACTERISTICS: MARSTERS, JR., JAMES C.
; ADDRESSEE: THE IDENTIFICATION,
; ADDRESSEE: CHARACTERIZATION AND
; ADDRESSEE: INHIBITION OF
; ADDRESSEE: FARNESYLTRANSFERASE
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK/ASKII
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08062
; FILING DATE: AUGUST 24, 1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/935,087
; FILING DATE: 24 AUGUST 1992 (24.08.92)
; NAME: UNKNOWN
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD377PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577

```

```

; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US93-08062-16

Query Match 50.6%; Score 41; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TQSPQC 15
Db 1 TQSPQC 7

RESULT 13
US-09-270-767-45686
; Sequence 45686, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45686
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-45686

Query Match 49.4%; Score 40; DB 4; Length 136;
Best Local Similarity 46.7%; Pred. No. 39;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQC 15
Db 11 LLTNTDPRTRTESRC 25

RESULT 14
US-09-949-016-7849
; Sequence 7849, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7849
; LENGTH: 1031
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-7849

Query Match 49.4%; Score 40; DB 4; Length 1031;

```

```

Best Local Similarity 53.8%; Pred. No. 3.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQ 13
Db 225 LTGESEPQTRSPF 237

RESULT 15
US-09-252-991A-25786
; Sequence 25786, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25786
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25786

Query Match 48.1%; Score 39; DB 4; Length 144;
Best Local Similarity 53.8%; Pred. No. 60;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPQNC 15
Db 96 GSTSPTTASAPNC 108

RESULT 16
US-07-800-364B-14
; Sequence 14, Application US/07800364B
; Patent No. 5688578
; GENERAL INFORMATION:
; APPLICANT: Hewick, Rodney M.
; APPLICANT: Wang, Jack H.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/800,364B
; FILING DATE: 26-NOV-1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI 5182A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-5851
; TELEFAX: 617-876-5851

```

```

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-800-364B-14

Query Match 48.1%; Score 39; DB 1; Length 281;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12
Db 114 LLGQAPRSQOP 125

RESULT 17
US-07-989-847-12
; Sequence 12, Application US/07989847
; Patent No. 5866364
; GENERAL INFORMATION:
; APPLICANT: Israel, David
; APPLICANT: Wolfman, Neil M.
; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein
; TITLE OF INVENTION: Heterodimers, Compositions and Methods of Use.
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140-2387
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,847
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI-5192B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-989-847-12

Query Match 48.1%; Score 39; DB 2; Length 281;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12
Db 114 LLGQAPRSQOP 125

RESULT 18
US-08-469-411-12
; Sequence 12, Application US/08469411
; Patent No. 6190880
; GENERAL INFORMATION:
; APPLICANT: Israel, David

```

;; Wolfman, Neil M.  
;; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein  
;; NUMBER OF SEQUENCES: 30  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.  
;; STREET: 87 CambridgePark Drive  
;; CITY: Cambridge  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02140-2387  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Tape  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/469,411  
;; FILING DATE: 06-Jun-1995  
;; CLASSIFICATION: <Unknown>  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kapinos, Ellen J.  
;; REGISTRATION NUMBER: 32,245  
;; REFERENCE/DOCKET NUMBER: GI-5192B-CON  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-498-8622  
;; TELEFAX: 617-876-5851  
;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 281 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-469-411-12  
;;  
;; Query Match 48.1%; Score 39; DB 3; Length 281;  
;; Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
;; Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
;;  
Qy 1 LLGNSSPRTQSP 12  
Db 114 LLGQAPRSQSP 125  
;;  
RESULT 19  
US-09-780-601A-12  
;; Sequence 12, Application US/09780601A  
;; Patent No. 6593109  
;; GENERAL INFORMATION:  
;; APPLICANT: Israel, David  
;; APPLICANT: Wolfman, Neil M.  
;; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein  
;; NUMBER OF SEQUENCES: 30  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.  
;; STREET: 87 CambridgePark Drive  
;; CITY: Cambridge  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02140-2387  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Tape  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/780,601A  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/469,411

;; FILING DATE: 06-Jun-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kapinos, Ellen J.  
;; REGISTRATION NUMBER: 32,245  
;; REFERENCE/DOCKET NUMBER: GI-5192B-CON  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-498-8622  
;; TELEFAX: 617-876-5851  
;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 281 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-09-780-601A-12  
;;  
;; Query Match 48.1%; Score 39; DB 4; Length 281;  
;; Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
;; Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
;;  
Qy 1 LLGNSSPRTQSP 12  
Db 114 LLGQAPRSQSP 125  
;;  
RESULT 20  
PCT-US91-07635-6  
;; Sequence 6, Application PC/TUS9107635  
;; GENERAL INFORMATION:  
;; APPLICANT: OPPERMANN, HERMANN  
;; APPLICANT: OZKAYNAK, ENGIN  
;; APPLICANT: RUEGER, DAVID C  
;; APPLICANT: KUBERASAMPATH, THANGAVEL  
;; TITLE OF INVENTION: OSTEOGENIC DEVICES  
;; NUMBER OF SEQUENCES: 9  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: TESTA, HURWITZ & THIBRAULT  
;; STREET: 53 STATE STREET  
;; CITY: BOSTON  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US91/07635  
;; FILING DATE: 19911018  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: PITCHER ESQ, EDMUND R  
;; REGISTRATION NUMBER: 27,829  
;; REFERENCE/DOCKET NUMBER: CRP-056PC  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617/248-7000  
;; TELEFAX: 617/248-7100  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 399 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; PCT-US91-07635-6  
;;  
;; Query Match 48.1%; Score 39; DB 5; Length 399;  
;; Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
;; Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
;;  
Qy 1 LLGNSSPRTQSP 12  
Db 232 LLGQAPRSQSP 243

RESULT 21  
US-07-841-646-29  
; Sequence 29, Application US/07841646  
; Patent No. 5266583  
; GENERAL INFORMATION:  
; APPLICANT: OPPERMANN, HERMANN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H.L.  
; TITLE OF INVENTION: OSTEOGENIC DEVICES  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: U.S.A.  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/841.646  
; FILING DATE: 19920221  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 810,560  
; FILING DATE: 20-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 827,052  
; FILING DATE: 28-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 660,162  
; FILING DATE: 22-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 621,988  
; FILING DATE: 04-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 621,849  
; FILING DATE: 04-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 616,374  
; FILING DATE: 21-NOV-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 600,024  
; FILING DATE: 18-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 599,543  
; FILING DATE: 18-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 579,865  
; FILING DATE: 07-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 569,920  
; FILING DATE: 20-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 483,913  
; FILING DATE: 22-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 422,613  
; FILING DATE: 17-OCT-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 315,342  
; FILING DATE: 23-FEB-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 232,630  
; FILING DATE: 15-AUG-1988  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 179,460  
; FILING DATE: 08-APR-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PITCHER, EDMUND R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-001CP6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/248-7000  
; TELEFAX: 617/248-7100  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-841-646-29  
  
Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 LLGNSPRTQSP 12  
Db 235 LLGQAPRSQOP 246  
||| :|||  
||| :|||  
  
RESULT 22  
US-07-901-703-11  
; Sequence 11, Application US/07901703  
; Patent No. 5344654  
; GENERAL INFORMATION:  
; APPLICANT: RUEGER, DAVID C  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: OPPERMANN, HERMANN  
; APPLICANT: OZKAYNAK, ENGIN  
; TITLE OF INVENTION: PROSTHETIC DEVICES HAVING ENHANCED  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: EXCHANGE PLACE, 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/901,703  
; FILING DATE: 19920616  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PITCHER, ESQ, EDMUND R  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: STK-057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/248-7000  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-901-703-11  
  
Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 LLGNSPRTQSP 12

Db 235 LLQAPRSQP 246  
||| :||:|  
FILING DATE: 15-AUG-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 179,460  
FILING DATE: 08-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER, EDMUND R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-001CP6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/248-7000  
TELEFAX: 617/248-7100  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-147-023-29

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLQAPRSQP 246

RESULT 24  
US-08-206-864-4  
Sequence 4, Application US/08206864  
Patent No. 5610021  
GENERAL INFORMATION:  
APPLICANT: RUEGER, DAVID C  
APPLICANT: JONES, WILLIAM K  
APPLICANT: TUCKER, RONALD F  
APPLICANT: OPPERMAN, HERMANN  
APPLICANT: OZKAYNAK, ENGIN  
APPLICANT: KUBERASAMPATH, THANGAVEL  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR  
RECOMBINANT OSTEOGENIC PROTEIN PRODUCTION  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,  
ADDRESSEE: INC.  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/206,864  
FILING DATE: 04-MAR-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/027,070  
FILING DATE: 04-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/841,646  
FILING DATE: 21-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER, EDMUND R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-096  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/7000  
TELEFAX: 617/248-7100  
INFORMATION FOR SEQ ID NO: 4:

Db 235 LLQAPRSQP 246  
||| :||:|  
FILING DATE: 15-AUG-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 179,460  
FILING DATE: 08-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER, EDMUND R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-001CP6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/248-7000  
TELEFAX: 617/248-7100  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-147-023-29

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLQAPRSQP 246

RESULT 24  
US-08-206-864-4  
Sequence 4, Application US/08206864  
Patent No. 5610021  
GENERAL INFORMATION:  
APPLICANT: RUEGER, DAVID C  
APPLICANT: JONES, WILLIAM K  
APPLICANT: TUCKER, RONALD F  
APPLICANT: OPPERMAN, HERMANN  
APPLICANT: OZKAYNAK, ENGIN  
APPLICANT: KUBERASAMPATH, THANGAVEL  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR  
RECOMBINANT OSTEOGENIC PROTEIN PRODUCTION  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,  
ADDRESSEE: INC.  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/206,864  
FILING DATE: 04-MAR-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/027,070  
FILING DATE: 04-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/841,646  
FILING DATE: 21-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER, EDMUND R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-096  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/7000  
TELEFAX: 617/248-7100  
INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-206-864-4

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQSP 246

## RESULT 25

US-08-278-729A-21  
; Sequence 21, Application US/08278729A  
; Patent No. 5650276

## GENERAL INFORMATION:

; APPLICANT: SMART, JOHN  
; APPLICANT: OPPERMAN, HERMAN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H.L.  
; APPLICANT: COHEN, CHARLES M.

; TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
; STREET: 45 SOUTH STREET  
; CITY: HOPKINTON  
; STATE: MA

; COUNTRY: USA

; ZIP: 01748

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/278,729A

; FILING DATE: 20-JUL-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: PITCHER ESQ., EDMUND R.

; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: CRP-058CPFW

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (508) 435-9001

; TELEFAX: (508) 435-6951

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 402 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-278-729A-21

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQSP 246

## RESULT 26

US-08-480-528A-8

; Sequence 8, Application US/08480528A

; Patent No. 5652118  
; GENERAL INFORMATION:  
; APPLICANT: OPPERMAN, HERMANN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H.L.  
; APPLICANT: COHEN, CHARLES M.

; TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
; STREET: 45 SOUTH STREET  
; CITY: HOPKINTON  
; STATE: MA

; COUNTRY: USA

; ZIP: 01748

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/480,528A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: FENTON ESQ., GILLIAN M.

; REGISTRATION NUMBER: 36,508

; REFERENCE/DOCKET NUMBER: CRP-076FW

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 248-7560

; TELEFAX: (617) 248-7100

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 402 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-480-528A-8

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQSP 246

## RESULT 27

US-08-479-666-8

; Sequence 8, Application US/08479666

; Patent No. 5652337

; GENERAL INFORMATION:

; APPLICANT: OPPERMAN, HERMANN

; APPLICANT: OZKAYNAK, ENGIN

; APPLICANT: KUBERASAMPATH, THANGAVEL

; APPLICANT: RUEGER, DAVID C.

; APPLICANT: PANG, ROY H.L.

; APPLICANT: COHEN, CHARLES M.

; TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
; STREET: 45 SOUTH STREET  
; CITY: HOPKINTON  
; STATE: MA

; COUNTRY: USA

; ZIP: 01748

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,666  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: FENTON Esq., GILLIAN M.  
REGISTRATION NUMBER: 36,508  
REFERENCE/DOCKET NUMBER: CRP-076DV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 435-9001  
TELEFAX: (508) 435-6951  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-479-666-8

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12  
||| :||:|  
Db 235 LLGORAPRSQQP 246

RESULT 28  
US-08-155-343A-21  
Sequence 21, Application US/08155343A  
Patent No. 5656593  
GENERAL INFORMATION:  
APPLICANT: KUBERASAMPATH, THANGAVEL  
APPLICANT: RUEGER, DAVID C.  
APPLICANT: OPPERMANN, HERMAN  
APPLICANT: COHEN, CHARLES M.  
APPLICANT: PANG, ROY H.L.  
TITLE OF INVENTION: MORPHOGENIC-INDUCED PERIODONTAL TISSUE  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/155,343A  
FILING DATE: 15-NOV-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: FENTON Esq., GILLIAN M.  
REGISTRATION NUMBER: 36,508  
REFERENCE/DOCKET NUMBER: CRP-067FW  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7560  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-155-343A-21

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12  
||| :||:|  
Db 235 LLGORAPRSQQP 246

RESULT 29  
US-08-406-672-21  
Sequence 21, Application US/08406672  
Patent No. 56744844  
GENERAL INFORMATION:  
APPLICANT: KUBERASAMPATH, THANGAVEL  
APPLICANT: COHEN, CHARLES M.  
APPLICANT: OPPERMANN, HERMAN  
APPLICANT: OZKAYNAK, ENGIN  
APPLICANT: RUEGER, DAVID C.  
APPLICANT: PANG, ROY H.L.  
TITLE OF INVENTION: TREATMENT TO PREVENT LOSS OF AND/OR  
TITLE OF INVENTION: INCREASE BONE MASS IN METABOLIC BONE DISEASES  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/406,672  
FILING DATE: 20-MAR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 752,857  
FILING DATE: 30-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 667,274  
FILING DATE: 11-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: FENTON Esq., GILLIAN M.  
REGISTRATION NUMBER: 36,508  
REFERENCE/DOCKET NUMBER: CRP-060CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7560  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-406-672-21

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12  
||| :||:|  
Db 235 LLGORAPRSQQP 246

RESULT 30  
US-08-643-563A-21  
Sequence 21, Application US/08643563A

Patent No. 5707810  
; GENERAL INFORMATION:  
; APPLICANT: SMART, JOHN  
; APPLICANT: OPPERMANN, HERMAN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H. L.  
; APPLICANT: COHEN, CHARLES M.  
; TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
; STREET: 45 SOUTH STREET  
; CITY: HOPKINTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 01748  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/643,563A  
; FILING DATE: 06-MAY-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: TWOMEY Esq., MICHAEL J.  
; REGISTRATION NUMBER: 38,349  
; REFERENCE/DOCKET NUMBER: CRP-058CN2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508) 435-9001  
; TELEFAX: (508) 435-6951  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-643-563A-21

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQOP 246

RESULT 31  
US-08-447-570-29  
; Sequence 29, Application US/08447570  
; Patent No. 5714589  
; GENERAL INFORMATION:  
; APPLICANT: OPPERMANN, HERMAN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H. L.  
; TITLE OF INVENTION: OSTEOGENIC DEVICES  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: U.S.A.  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,570  
; FILING DATE: 21-FEB-1992  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 810,560  
; FILING DATE: 20-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 827,052  
; FILING DATE: 28-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 660,162  
; FILING DATE: 22-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 621,988  
; FILING DATE: 04-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 621,849  
; FILING DATE: 04-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 616,374  
; FILING DATE: 21-NOV-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 600,024  
; FILING DATE: 18-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 599,543  
; FILING DATE: 18-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 579,865  
; FILING DATE: 07-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 569,920  
; FILING DATE: 20-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 483,913  
; FILING DATE: 22-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 422,613  
; FILING DATE: 17-OCT-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 315,342  
; FILING DATE: 23-FEB-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 232,630  
; FILING DATE: 15-AUG-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 179,460  
; FILING DATE: 08-APR-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PITCHER, EDMUND R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-001CP6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/248-7000  
; TELEFAX: 617/248-7100  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-447-570-29

Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQOP 246

RESULT 32  
US-08-643-763A-21  
; Sequence 21, Application US/08643763A  
; Patent No. 5733878  
; GENERAL INFORMATION:  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: OPPERMANN, HERMAN  
; APPLICANT: COHEN, CHARLES M.  
; APPLICANT: PANG, ROY H.L.  
; TITLE OF INVENTION: MORPHOGENIC-INDUCED PERIODONTAL TISSUE  
; TITLE OF INVENTION: REGENERATION.  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
; STREET: 45 SOUTH STREET  
; CITY: HOPKINTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 01748  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/643,763A  
; FILING DATE: 06-MAY-1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FENTON Esq., GILLIAN M.  
; REGISTRATION NUMBER: 36,508  
; REFERENCE/DOCKET NUMBER: CRP-067CN  
; TELEPHONE: (617) 248-7560  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-643-763A-21  
Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQQP 246  
RESULT 33  
US-08-462-623-21  
; Sequence 21, Application US/08462623  
; Patent No. 5739107  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, CHARLES M.  
; APPLICANT: CHARETTE, MARC F.  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: OPPERMANN, HERMAN  
; APPLICANT: PANG, ROY H.L.  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: SMART, JOHN E.  
; TITLE OF INVENTION: MORPHOGEN TREATMENT OF GASTROINTESTINAL  
; TITLE OF INVENTION: ULCERS.  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES

STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,623  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/445,882  
FILING DATE: 22-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: FENTON Esq., GILLIAN M.  
REGISTRATION NUMBER: 36,508  
REFERENCE/DOCKET NUMBER: CRP-074CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 435-9001  
TELEFAX: (508) 435-6951  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-462-623-21  
Query Match 48.1%; Score 39; DB 1; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQQP 246  
RESULT 34  
US-08-451-953A-21  
; Sequence 21, Application US/08451953A  
; Patent No. 5741641  
; GENERAL INFORMATION:  
; APPLICANT: SMART, JOHN  
; APPLICANT: OPPERMANN, HERMAN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H.L.  
; APPLICANT: COHEN, CHARLES M.  
; TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
; STREET: 45 SOUTH STREET  
; CITY: HOPKINTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 01748  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/451,953A  
FILING DATE: 26-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER Esq., EDMUND R.

```

; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-058CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-451-953A-21

Query Match 48.1%; Score 39; DB 1; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12
Db 235 LLGQAPRSQOP 246

RESULT 35
US-08-459-346-6
; Sequence 6, Application US/08459346
; Patent No. 5834179
; GENERAL INFORMATION:
; APPLICANT: JONES, WILLIAM K
; APPLICANT: TUCKER, RONALD F
; APPLICANT: RUEGER, DAVID C
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
; INC.
; STREET: 35 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,346
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/029,335
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,091
; FILING DATE: 03-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,235
; FILING DATE: 16-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,336
; FILING DATE: 08-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,780
; FILING DATE: 31-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-081CP
; INFORMATION FOR SEQ ID NO: 6:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-346-6

Query Match 48.1%; Score 39; DB 2; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12
Db 235 LLGQAPRSQOP 246

RESULT 36
US-08-445-468A-21
; Sequence 21, Application US/08445468A
; Patent No. 5849686
; GENERAL INFORMATION:
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: OPPERMANN, HERMAN
; APPLICANT: PANG, ROY H.L.
; APPLICANT: COHEN, CHARLES M.
; TITLE OF INVENTION: MORPHOGEN-INDUCED LIVER REGENERATION
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,468A
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON ESQ., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-072FW2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7560
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-445-468A-21

Query Match 48.1%; Score 39; DB 2; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12
Db 235 LLGQAPRSQOP 246

RESULT 37
US-08-901-200A-8
; Sequence 8, Application US/08901200A
; Patent No. 5854071
; GENERAL INFORMATION:

```

APPLICANT: OPPERMANN, HERMANN  
APPLICANT: OZKAYNAK, ENGIN  
APPLICANT: KUBERASAMPATH, THANGAVEL  
APPLICANT: RUEGER, DAVID C.  
APPLICANT: PANG, ROY H.L.  
APPLICANT: COHEN, CHARLES M.  
TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/901,200A  
FILING DATE: 28-JUL-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: MEYERS, THOMAS C.  
REGISTRATION NUMBER: 36,989  
REFERENCE/DOCKET NUMBER: CRP-076DV2  
TELEPHONE: (508) 435-9001  
TELEFAX: (508) 435-6951  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-901-200A-8

Query Match 48.1%; Score 39; DB 2; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LIGNSPRTQSP 12  
|||:|:|:  
Db 235 LIGORAPRSQP 246

RESULT 38  
US-08-449-700-29 Application US/08449700  
Sequence 29, Appication US/08449700  
Patent No. 5863758  
GENERAL INFORMATION:  
APPLICANT: OPPERMANN, HERMANN  
APPLICANT: OZKAYNAK, ENGIN  
APPLICANT: KUBERASAMPATH, THANGAVEL  
APPLICANT: RUEGER, DAVID C.  
APPLICANT: PANG, ROY H.L.  
TITLE OF INVENTION: OSTEOGENIC DEVICES  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
STREET: 53 STATE STREET  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/449,700  
FILING DATE: 21-FEB-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 810,560  
FILING DATE: 20-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 827,052  
FILING DATE: 28-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 660,162  
FILING DATE: 22-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 621,988  
FILING DATE: 04-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 621,849  
FILING DATE: 04-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 616,374  
FILING DATE: 21-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 600,024  
FILING DATE: 18-OCT-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 599,543  
FILING DATE: 18-OCT-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 579,865  
FILING DATE: 07-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 569,920  
FILING DATE: 20-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 483,913  
FILING DATE: 22-FEB-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 422,613  
FILING DATE: 17-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 315,342  
FILING DATE: 23-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 232,630  
FILING DATE: 15-AUG-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 179,460  
FILING DATE: 08-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: FITCHER, EDMUND R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-001CP6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/248-7000  
TELEFAX: 617/248-7100  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-449-700-29

Query Match 48.1%; Score 39; DB 2; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LIGNSPRTQSP 12  
|||:|:|:  
Db 235 LIGORAPRSQP 246

RESULT 39

US-08-449-699A-29  
; Sequence 29, Application US/08449699A  
; Patent No. 5958441  
; GENERAL INFORMATION:  
; APPLICANT: OPPERMANN, HERMANN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H.L.  
; TITLE OF INVENTION: ANTIBODIES TO OSTEOGENIC PROTEINS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: 125 HIGH STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: U.S.A.  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/449,699A  
; FILING DATE: 24-MAY-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/147,023  
; FILING DATE: 01-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PITCHER, EDMUND R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: STK-001CP6CN  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/248-7000  
; TELEFAX: 617/248-7100  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-449-699A-29  
Query Match 48.1%; Score 39; DB 2; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLGORAPRSQQP 246  
RESULT 40  
US-08-461-397A-21  
; Sequence 21, Application US/08461397A  
; Patent No. 5972884  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, CHARLES M.  
; APPLICANT: CHARETTE, MARC F.  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: OPPERMANN, HERMANN  
; APPLICANT: PANG, ROY H.L.  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: SMART, JOHN E.  
; TITLE OF INVENTION: MORPHOGEN TREATMENT FOR LIMITING  
; PROLIFERATION OF EPITHELIAL CELLS.  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
; STREET: 45 SOUTH STREET

CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,397A  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PITCHER ESQ, EDMUND R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-074FW2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/248-7000  
; TELEFAX: 617/248-7100  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 402 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-397A-21  
Query Match 48.1%; Score 39; DB 2; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLGORAPRSQQP 246  
RESULT 41  
US-08-912-088-21  
; Sequence 21, Application US/08912088  
; Patent No. 5994131  
; GENERAL INFORMATION:  
; APPLICANT: SMART, JOHN  
; APPLICANT: OPPERMANN, HERMANN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H.L.  
; APPLICANT: COHEN, CHARLES M.  
; TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
; STREET: 45 SOUTH STREET  
; CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/912,088  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,729  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PITCHER ESQ., EDMUND R.  
; REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-058CPFW  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 435-9001  
TELEFAX: (508) 435-6951  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-912-088-21

Query Match 48.1%; Score 39; DB 2; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQP 246

RESULT 42  
US-08-278-730A-21  
Sequence 21, Application US/08278730A  
Patent No. 6022853  
GENERAL INFORMATION:  
APPLICANT: KUBERASAMPATH, THANGAVEL  
APPLICANT: COHEN, CHARLES M.  
APPLICANT: RUEGER, DAVID C.  
APPLICANT: OPPERMAN, HERMAN  
APPLICANT: PANG, ROY H.L.  
TITLE OF INVENTION: MORPHOGENIC-ENRICHED DIETARY COMPOSITION  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA USA  
COUNTRY: USA  
ZIP: 01748

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/278,730A  
FILING DATE: 20-JULY-1994  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: FENTON Esq., GILLIAN M.  
REGISTRATION NUMBER: 36,508  
REFERENCE/DOCKET NUMBER: CRP-071FW  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7560  
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-278-730A-21

Query Match 48.1%; Score 39; DB 3; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQP 246

RESULT 43  
US-08-889-419-6  
Sequence 6, Application US/08889419  
Patent No. 6071708  
GENERAL INFORMATION:  
APPLICANT: JONES, WILLIAM K  
APPLICANT: TUCKER, RONALD F  
APPLICANT: RUEGER, DAVID C  
APPLICANT: OPPERMAN, HERMANN  
APPLICANT: OZKAYNAK, ENGIN  
APPLICANT: KUBERASAMPATH, THANGAVEL  
TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patent Administrator, Testa, Hurwitz &  
ADDRESSEE: Thibeault, LLP  
STREET: 125 HIGH STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/889,419  
FILING DATE: 08-JUL-1997  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/459,346  
FILING DATE: 02-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MEYERS, THOMAS C.  
REGISTRATION NUMBER: 36,989  
REFERENCE/DOCKET NUMBER: CRP-081DVCN

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-889-419-6

Query Match 48.1%; Score 39; DB 3; Length 402;  
Best Local Similarity 58.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSPRTQSP 12  
||| :||:|  
Db 235 LLGQAPRSQP 246

RESULT 44  
US-08-445-467-21  
Sequence 21, Application US/08445467  
Patent No. 6077823  
GENERAL INFORMATION:  
APPLICANT: KUBERASAMPATH, THANGAVEL  
APPLICANT: PANG, ROY HL  
APPLICANT: OPPERMAN, HERMANN  
APPLICANT: RUEGER, DAVID C  
APPLICANT: COHEN, CHARLES M  
APPLICANT: OZKAYNAK, ENGIN  
APPLICANT: SMART, JOHN E  
TITLE OF INVENTION: MORPHOGEN-INDUCED MODULATION OF  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CREATIVE BIOMOLECULES, INC.  
STREET: 35 SOUTH STREET

```
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,467
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/165,511
; FILING DATE:
; APPLICATION NUMBER: US/07/938,336
; FILING DATE:
; APPLICATION NUMBER: US/07/667,274
; FILING DATE: 11-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/753,059
; FILING DATE: 30-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,764
; FILING DATE: 30-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-059CP.APP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-445-467-21

Query Match 48.1%; Score 39; DB 3; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12
Db 235 LLGQAPRSQQP 246

RESULT 45
US-08-480-515A-21
; Sequence 21, Application US/08480515A
; Patent No. 6090776
; GENERAL INFORMATION:
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: PANG, ROY H.L.
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: COHEN, CHARLES M.
; APPLICANT: SMART, JOHN E.
; TITLE OF INVENTION: MORPHOGEN TREATMENT OF ORGAN TRANSPLANTS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,467
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/901,200
; FILING DATE: 28-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, THOMAS C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: CRP-076DV2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
; US-08-445-467-21

Query Match 48.1%; Score 39; DB 3; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12
Db 235 LLGQAPRSQQP 246

RESULT 46
US-09-219-391-8
; Sequence 8, Application US/09219391
; Patent No. 6153583
; GENERAL INFORMATION:
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; APPLICANT: COHEN, CHARLES M.
; TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/219,391
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/901,200
; FILING DATE: 28-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, THOMAS C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: CRP-076DV2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
; US-08-480-515A-21

Query Match 48.1%; Score 39; DB 3; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSP 12
Db 235 LLGQAPRSQQP 246

RESULT 46
US-09-219-391-8
; Sequence 8, Application US/09219391
; Patent No. 6153583
; GENERAL INFORMATION:
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; APPLICANT: COHEN, CHARLES M.
; TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/219,391
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/901,200
; FILING DATE: 28-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MEYERS, THOMAS C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: CRP-076DV2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
; US-08-480-515A-21
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-219-391-8

Query Match      48.1%; Score 39; DB 3; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSSPRQSP 12
   |||:|:|
Db 235 LLGQAPRSQQP 246

RESULT 47
US-09-170-936-21
; Sequence 21, Application US/09170936
; Patent No. 6333312
; GENERAL INFORMATION:
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: COHEN, CHARLES M.
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; APPLICANT: SMART, JOHN E.
; TITLE OF INVENTION: TREATMENT TO PREVENT LOSS OF AND/OR
; TITLE OF INVENTION: INCREASE BONE MASS IN METABOLIC BONE DISEASES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09170,936
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,883
; FILING DATE: 2-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON Esq., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-060CPFWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7560
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-170-936-21

Query Match      48.1%; Score 39; DB 3; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSSPRQSP 12
   |||:|:|
Db 235 LLGQAPRSQQP 246

RESULT 48
US-08-402-542-6
; Sequence 6, Application US/08402542
; Patent No. 6395883
; GENERAL INFORMATION:
; APPLICANT: JONES, WILLIAM K
; APPLICANT: TUCKER, RONALD F
; APPLICANT: RUEGER, DAVID C
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
; TITLE OF INVENTION: OF MATTER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
; INC.
; STREET: 35 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,542
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/040,510
; FILING DATE:
; APPLICATION NUMBER: US 08/029,335
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,091
; FILING DATE: 03-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,235
; FILING DATE: 16-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,336
; FILING DATE: 08-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,780
; FILING DATE: 31-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-081CP
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-402-542-6

Query Match      48.1%; Score 39; DB 3; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSSPRQSP 12
   |||:|:|
Db 235 LLGQAPRSQQP 246

RESULT 49
US-08-461-113-21
; Sequence 21, Application US/08461113
; Patent No. 6399569
; GENERAL INFORMATION:
; APPLICANT: COHEN, CHARLES M.
```

```

; APPLICANT: CHARETTE, MARC F.
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: PANG, ROY H.L.
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: SMART, JOHN E.
; TITLE OF INVENTION: MORPHOGEN TREATMENT FOR LIMITING
; TITLE OF INVENTION: PROLIFERATION OF EPITHELIAL CELLS.
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,113
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/445,882
; FILING DATE: 22-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON Esq., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-074DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-113-21

```

```

Query Match 48.1%; Score 39; DB 3; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 LLGNSSPRTQSP 12
Db 235 LLGQAPRSQQP 246

```

```

RESULT 50
US-08-456-033-21
; Sequence 21, Application US/08456033
; Patent No. 6495513
; GENERAL INFORMATION:
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: PANG, ROY H.L.
; APPLICANT: COHEN, CHARLES M.
; TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND
; TITLE OF INVENTION: REPAIR.
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA

```

```

; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,033
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON Esq., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-070DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-456-033-21

```

```

Query Match 48.1%; Score 39; DB 4; Length 402;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 LLGNSSPRTQSP 12
Db 235 LLGQAPRSQQP 246

```

```

Search completed: August 25, 2005, 11:45:23
Job time : 44 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 11:34:33 ; Search time 40 Seconds  
(without alignments)  
36.081 Million cell updates/sec

Title: US-10-803-541-2

Perfect score: 81

Sequence: 1 LLGSSPRTQSPQNC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR\_79:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	664	1	VEHULA
2	73	90.1	665	2	S27267
3	69	85.2	665	2	S28182
4	44	54.3	521	2	F87775
5	44	54.3	970	2	S63059
6	42	51.9	291	2	E97463
7	42	51.9	690	2	S28222
8	42	51.9	2150	2	S71629
9	41	50.6	148	2	AH2488
10	41	50.6	439	2	A12098
11	41	50.6	542	1	T02074
12	41	50.6	542	2	T03789
13	41	50.6	599	2	I53395
14	41	50.6	717	2	S31034
15	41	50.6	717	2	S31035
16	41	50.6	2038	2	A43742
17	40.5	50.0	470	2	A71466
18	40	49.4	94	2	E82515
19	40	49.4	199	2	S56155
20	40	49.4	212	2	H90369
21	40	49.4	454	2	I51916
22	40	49.4	551	2	A12964
23	40	49.4	690	2	D98318
24	40	49.4	1017	2	A37227
25	40	49.4	1020	2	A34474
26	40	49.4	1020	2	B24639
27	40	49.4	1031	2	I51674
28	40	49.4	1033	2	I49143
29	40	49.4	1033	2	A25344

30	40	49.4	1034	2	JN0903	H+/K+-exchanging A
31	40	49.4	1034	2	A31671	H+/K+-exchanging A
32	40	49.4	1035	2	A35292	H+/K+-exchanging A
33	40	49.4	1035	2	S23406	H+/K+-exchanging A
34	39	48.1	169	2	S20484	hypothetical prote
35	39	48.1	402	2	A45056	osteogenic protein
36	39	48.1	403	2	B46165	envelope surface g
37	39	48.1	408	2	C46165	hypothetical prote
38	39	48.1	481	2	T16484	CED-6 protein - Ca
39	39	48.1	492	2	T43064	coat protein gamma
40	39	48.1	563	2	S33957	env polyprotein pr
41	39	48.1	642	1	VCWVS2	env polyprotein pr
42	39	48.1	642	2	T10533	env polyprotein pr
43	39	48.1	645	1	VCWVS1	lamin - chicken
44	39	48.1	657	2	S05517	env polyprotein -
45	39	48.1	662	1	VCWVGF	env polyprotein -
46	39	48.1	662	2	A25982	transcription regu
47	39	48.1	727	2	B60191	filensin - bovine
48	39	48.1	755	2	S32103	EVI1 protein - hum
49	39	48.1	1042	2	S41705	transcription regu
50	39	48.1	1042	2	A31591	oncogene Evi-1 - h
51	39	48.1	1051	2	A60191	nicotinic acetylch
52	38.5	47.5	470	2	A39218	KIAA1424 protein [
53	38.5	47.5	1944	2	A59438	AUR protein - huma
54	38.5	47.5	5262	2	T03454	hypothetical prote
55	38	46.9	81	2	AH2564	Na+/K+-exchanging
56	38	46.9	220	2	S20970	probable transcrip
57	38	46.9	227	2	D83543	hypothetical prote
58	38	46.9	278	2	F85012	conserved hypother
59	38	46.9	289	2	T42994	hypothetical prote
60	38	46.9	290	2	T44550	conserved hypother
61	38	46.9	292	2	T40045	hypothetical prote
62	38	46.9	305	2	T45062	hypothetical prote
63	38	46.9	353	2	S22805	env polyprotein -
64	38	46.9	381	2	T10707	adenosylmethionine
65	38	46.9	445	1	VCFMVL	env polyprotein -
66	38	46.9	522	2	E84833	hypothetical prote
67	38	46.9	534	1	VCWVSF	env polyprotein -
68	38	46.9	575	2	T01552	hypothetical prote
69	38	46.9	627	1	VCVMW2	env polyprotein -
70	38	46.9	630	2	A39344	tumor-associated m
71	38	46.9	631	2	I52257	episialin - mouse
72	38	46.9	636	1	VCWVFS	env polyprotein pr
73	38	46.9	639	1	VCWVSA	env polyprotein -
74	38	46.9	640	1	VCVMW1	env polyprotein pr
75	38	46.9	640	1	VCMVVR	gp70 protein - mur
76	38	46.9	644	2	S15464	env polyprotein -
77	38	46.9	662	1	VCMWLB	env polyprotein pr
78	38	46.9	665	1	VCMWKA	env polyprotein pr
79	38	46.9	665	1	VCMVVR	env polyprotein -
80	38	46.9	665	1	VCWVEM	env polyprotein pr
81	38	46.9	668	1	VCMVFP	env polyprotein -
82	38	46.9	669	1	VCWVEK	envelope protein -
83	38	46.9	669	2	A46511	env polyprotein pr
84	38	46.9	676	1	VCMPVP	env protein - muri
85	38	46.9	676	2	T01381	env polyprotein -
86	38	46.9	676	2	S70395	env polyprotein -
87	38	46.9	688	2	A43491	env polyprotein -
88	38	46.9	689	2	B43491	env polyprotein -
89	38	46.9	781	2	A26641	Na+/K+-exchanging
90	38	46.9	821	2	S67087	hypothetical prote
91	38	46.9	899	2	F88391	protein R06B10.4 [
92	38	46.9	1010	2	B37227	Na+/K+-exchanging
93	38	46.9	1013	1	S00801	Na+/K+-exchanging
94	38	46.9	1013	2	C24639	Na+/K+-exchanging
95	38	46.9	1021	1	PWSHNA	Na+/K+-exchanging
96	38	46.9	1021	1	S04630	Na+/K+-exchanging
97	38	46.9	1021	2	A28199	Na+/K+-exchanging
98	38	46.9	1021	2	B24862	Na+/K+-exchanging
99	38	46.9	1022	2	S49127	Na+/K+-exchanging
100	38	46.9	1023	1	A24639	Na+/K+-exchanging

## ALIGNMENTS

## RESULT 1

VEHULA  
lamin A - human  
N:Alternate names: 70kDa lamin  
C;Species: Homo sapiens (man)  
C;Date: 28-May-1986 #sequence revision 04-Dec-1986 #text\_change 09-Jul-2004  
C;Accession: A02961; B24249; C24249  
R;McKeon, F. D.; Kirschner, M. W.; Caput, D.  
Nature 319, 463-468, 1986  
A;Title: Homologies in both primary and secondary structure between nuclear envelope and  
A;Reference number: A02962; MUID:86118697; PMID:3453101  
A;Accession: A02961  
A;Molecule type: mRNA  
A;Cross-references: UNIPROT:P02545; GB:X03444; NID:g34227; PIDN:CAA27173.1; PID:g34228  
A;Note: this sequence has been corrected in reference A94121; an omitted nucleotide cause  
R;Fisher, D. Z.; Chaudhary, N.; Blobel, G.  
Proc. Natl. Acad. Sci. U.S.A. 83, 6450-6454, 1986  
A;Title: cDNA sequencing of nuclear lamins A and C reveals primary and secondary structure  
A;Reference number: A94121; MUID:86313596; PMID:3462705  
A;Note: parts of sequences from rat lamins, but not human, were determined by protein se  
A;Accession: B24249  
A;Molecule type: mRNA  
A;Residues: 537-664 <PIS>  
A;Note: sequence fragment shown in publication  
A;Accession: C24249  
A;Molecule type: mRNA  
A;Residues: 150-664 <PI2>  
A;Cross-references: GB:M13452; NID:g196838; PIDN:AAA36160.1; PID:g386856  
A;Note: submitted sequence extracted from GenBank  
C;Comment: Lamins A and C (see CIR:VEHULC) are products of alternative splicing of the s  
C;Comment: The lamins (A, B, and C) contains several alpha-helical domains capable of fo  
C;Comment: The association of lamins, dependent upon ionic interactions, is interrupted  
rates with lamin' dissociation; it does not reform until telophase, when the lamins are d  
C;Genetics:  
A;Gene: GDB:LMNA; LMN1  
A;Cross-references: GDB:132146; OMIM:150330  
A;Map position: lq21.2-lq21.3  
C;Function:  
C;Description: structural component of the nuclear lamina, a fibrous meshwork on the nuc  
C;Superfamily: cytoskeletal keratin  
C;Keywords: alternative splicing; coiled coil; lipoprotein; membrane protein; methylated  
F;2-661/Product: lamin A #status predicted <NAT>  
F;2-33/Domain: head <HED>  
F;34-388/Domain: rod <ROD>  
F;34-70/Region: coil 1A  
F;81-218/Region: coil 1B  
F;243-388/Region: coil 2  
F;325/Region: stutter  
F;389-661/Domain: tail <END>  
F;417-420/Region: nuclear location signal  
F;661/Binding site: farnesyl (Cys) (covalent) #status predicted  
F;661/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 100.0%; Score 81; DB 1; Length 664;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
|||||  
Db 647 LLGNSSPRTQSPQNC 661

## RESULT 2

S27267  
lamin A - rat  
N;Contains: nucleoside-triphosphatase (EC 3.6.1.15)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 22-Nov-1993 #sequence revision 27-Jan-1995 #text\_change 10-Dec-1999  
C;Accession: S27267; S30573; A60832; S41934

R;Ozaki, T.; Sakiyama, S.  
FEBS Lett. 312, 165-168, 1992  
A;Title: Lamin A gene expression is specifically suppressed in v-src-transformed cells.  
A;Reference number: S27267; MUID:93050186; PMID:1426247  
A;Accession: S27267  
A;Molecule type: mRNA  
A;Residues: 1-665 <OZA>  
A;Cross-references: GB:X66870; NID:g56550; PIDN:CAA47342.1; PID:g56551  
A;Note: the authors translated the codon CAG for residue 258 as Glu  
A;Note: the nucleotide sequence in GB:X66870 (see reference S30573) differs from the pub.  
R;Ozaki, T.; Sakiyama, S.  
submitted to the EMBL Data Library, May 1992  
A;Reference number: S30573  
A;Accession: S30573  
A;Status: significant sequence differences  
A;Molecule type: mRNA  
A;Cross-references: GB:X66870; NID:g56550; PID:g56551  
A;Note: the nucleotide sequence in GB:X66870 differs from the published version (see ref  
R;Clawson, G. A.; Lackey, A.; Tokes, Z. A.  
Exp. Cell Res. 176, 180-186, 1988  
A;Title: The 46-kDa nucleoside triphosphatase of rat liver nuclear scaffold represents t  
A;Reference number: A60832; MUID:88225292; PMID:2836227  
A;Accession: A60832  
A;Molecule type: protein  
A;Residues: 12-25;29-33,'L',35-40,'K',42;198-202,'Q',204-209;282-296;352-366 <CLA>  
R;Jonnalagadda, V. S.; Parnaik, V. K.  
submitted to the EMBL Data Library, November 1993  
A;Reference number: S41934  
A;Accession: S41934  
A;Molecule type: mRNA  
A;Residues: 26-82,'A',84-469,'K',471-523,'T',525-583,'R',585,'R',587-605,'A',607-665 <JON  
A;Cross-references: EMBL:X76297; NID:g453179; PIDN:CAA53945.1; PID:g453180  
C;Superfamily: cytoskeletal keratin  
C;Keywords: blocked amino end; coiled coil; hydrolase; nuclear membrane; phosphoprotein  
F;1-390/Domain: head #status predicted <HEA>  
F;31-390/Domain: rod #status predicted <ROD>  
F;391-665/Domain: tail #status predicted <TAI>  
F;22,392,404,406/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 90.1%; Score 73; DB 2; Length 665;  
Best Local Similarity 93.3%; Pred. No. 0.00028;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
|||||  
Db 648 LLGNSSPRTQSPQNC 662

RESULT 3  
S28182  
lamin A - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 22-Nov-1993 #sequence revision 10-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S28182; S06662; S65931  
R;Nakajima, N.; Sado, T.  
Biochim. Biophys. Acta 1171, 311-314, 1993  
A;Title: Nucleotide sequence of a mouse lamin A cDNA and its deduced amino acid sequence  
A;Reference number: S28182; MUID:93144345; PMID:7916626  
A;Accession: S28182  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-665 <NAK>  
A;Cross-references: UNIPROT:Q91WF2; UNIPROT:Q9DC21  
R;Weber, K.; Flessmann, U.; Traub, P.  
FEBS Lett. 257, 411-414, 1989  
A;Title: Maturation of nuclear lamin A involves a specific carboxy-terminal trimming, wh  
a.  
A;Reference number: S06662; MUID:90060368; PMID:2583287  
A;Accession: S06662  
A;Molecule type: protein  
A;Residues: 520-580,'X',582-622,'V',624-647 <WEB>  
R;Nakajima, N.; Abe, K.  
FEBS Lett. 365, 108-114, 1995

A;Title: Genomic structure of the mouse A-type lamin gene locus encoding somatic and germline proteins  
A;Reference number: S65931; MUID:95300954; PMID:7781761  
A;Accession: S65931  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-117 <NAW>  
A;Cross-references: EMBL:D49732; NID:g1100724  
C;Superfamily: Cytoskeletal keratin  
C;Keywords: alternative splicing; coiled coil; intermediate filament; lipoprotein; nucleoside  
F;1-33/Domain: head #status predicted <HED>  
F;34-388/Domain: rod #status predicted <ROD>  
F;34-218/Region: coiled coil 1  
F;243-388/Region: coiled coil 2  
F;266/Region: heptad change of phase  
F;325/Region: stutter  
F;330/Region: heptad change of phase  
F;389-662/Domain: tail #status predicted <END>  
F;417-420/Region: nuclear location signal  
F;662/Binding site: farnesyl (Cys) (covalent) #status predicted

Query Match 85.2%; Score 69; DB 2; Length 665;  
Best Local Similarity 86.7%; Pred. No. 0.0013;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLGNSPRTQSPQC 15  
DB 648 LLGNSPRTQSPQC 662  
|||||:|||||

RESULT 4  
P87775  
Protein C24A11.8 [imported] - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C;Accession: F87775  
R;anonymouse, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes  
A;Reference number: AF5000; MUID:99069613; PMID:9851916  
A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 284, 2323-2328, 2001  
A;Accession: F87775  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-521 <STO>  
A;Cross-references: GB:chr\_I; PIDN:AAB37677.1; PID:g1703630; GSPDB:GN00019; CESP:C24A11.8  
C;Genetics:  
A;Gene: C24A11.8  
A;Map position: 1

Query Match 54.3%; Score 44; DB 2; Length 521;  
Best Local Similarity 61.5%; Pred. No. 16;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSPRTQSPQ 13  
DB 40 IIGNEPRTQSPQ 52  
|||||:|||||

RESULT 5  
S63059  
hypothetical protein YNL118c - yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein N1917  
C;Species: Saccharomyces cerevisiae  
C;Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 16-Aug-2004  
C;Accession: S63059; S59701; S67340  
R;De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.  
submitted to the Protein Sequence Database, April 1996  
A;Reference number: S63047  
A;Accession: S63059  
A;Molecule type: DNA  
A;Residues: 1-970 <DEA>  
A;Cross-references: UNIPROT:P53550; EMBL:Z71394; NID:g1302044; PID:e239800; PID:g1302045

A;Title: Genomic structure of the mouse A-type lamin gene locus encoding somatic and germline proteins  
A;Reference number: S65931; MUID:95300954; PMID:7781761  
A;Accession: S65931  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-117 <NAW>  
A;Cross-references: EMBL:D49732; NID:g1100724  
C;Superfamily: Cytoskeletal keratin  
C;Keywords: alternative splicing; coiled coil; intermediate filament; lipoprotein; nucleoside  
F;1-33/Domain: head #status predicted <HED>  
F;34-388/Domain: rod #status predicted <ROD>  
F;34-218/Region: coiled coil 1  
F;243-388/Region: coiled coil 2  
F;266/Region: heptad change of phase  
F;325/Region: stutter  
F;330/Region: heptad change of phase  
F;389-662/Domain: tail #status predicted <END>  
F;417-420/Region: nuclear location signal  
F;662/Binding site: farnesyl (Cys) (covalent) #status predicted

A;Experimental source: strain S288C  
R;Tzagoloff, A.A.  
submitted to the EMBL Data Library, June 1995  
A;Description: Suppressor of a yeast pet mutant.  
A;Reference number: S59701  
A;Accession: S59701  
A;Molecule type: DNA  
A;Residues: 1-424, 'L', 426-970 <TZA>  
A;Cross-references: EMBL:L43065; NID:g870733; PID:g870734  
A;Experimental source: strain D273-10B  
R;D'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.; Pandolfo, D.; Lanfranchi, G.  
submitted to the EMBL Data Library, February 1996  
A;Description: The DNA sequence of cosmid 14-13b from chromosome XIV of Saccharomyces cerevisiae  
A;Reference number: S67327  
A;Accession: S67340  
A;Molecule type: DNA  
A;Residues: 1-970 <DAN>  
A;Cross-references: EMBL:Z69382; NID:g1183941; PID:e221828; PID:g1183955  
C;Genetics:  
A;Gene: SGD:PSU1  
A;Cross-references: SGD:S0005062; MIPS:YNL118c  
A;Map position: 14L  
C;Superfamily: mutt domain homology  
F;129-163/Domain: mutt domain homology <MUTT>

Query Match 54.3%; Score 44; DB 2; Length 970;  
Best Local Similarity 75.0%; Pred. No. 29;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GNSPRTQSPON 14  
DB 595 GKSPSTQSKON 606  
|||||:|||||

RESULT 6  
E97463  
1564 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
C;Species: Agrobacterium tumefaciens  
C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: E97463  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, J.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: E97463  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-291 <KUR>  
A;Cross-references: UNIPROT:Q8UH30; GB:AE007869; PIDN:AAK86662.1; PID:g1515844; GSPDB:G1515844  
C;Genetics:  
A;Gene: AGR\_C1564  
A;Map position: circular chromosome

Query Match 51.9%; Score 42; DB 2; Length 291;  
Best Local Similarity 58.3%; Pred. No. 19;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 LGNSPRTQSPQ 13  
DB 18 LSHTTPRTQSPQ 29  
|||||:|||||

RESULT 7  
S28222  
peroxidase (EC 1.11.1.7) precursor - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: S28222; S25522  
R;Ng, S.W.; Wiedemann, M.; Kontermann, R.; Petersen, G.  
Biochim. Biophys. Acta 1171, 224-228, 1992  
A;Title: Molecular characterization of a putative peroxidase gene of Drosophila melanogaster  
A;Reference number: S28222; MUID:93129626; PMID:1482687

```
A;Accession: S28222
A;Molecule type: DNA
A;Residues: 1-690 <NS>
A;Cross-references: UNIPROT:Q01603; EMBL:X68131; NID:G2511639; PIDN:CAA48238.1; PID:G833
C;Genetics:
A;Gene: FlyBase:Pxd
A;Cross-references: FlyBase:FBgn0004577
A;Introns: 57/2; 126/2; 197/1; 445/2; 497/3; 621/1; 656/2
C;Superfamily: myeloperoxidase; myeloperoxidase homology
C;Keywords: glycoprotein; oxidoreductase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-690/Product: myeloperoxidase #status predicted <NAT>
F;310/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 51.9%; Score 42; DB 2; Length 690;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQN 14
::|||::|::|::|
Db 55 VIGNSLPRTSPAPQN 68

RESULT 8
S71629
sensory transduction histidine kinase dhka - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004
C;Accession: S71629
R;Wang, N.; Shauleky, G.; Escalante, R.; Loomis, W.F.
EMBO J. 15, 3890-3898, 1996
A;Title: A two-component histidine kinase gene that functions in Dictyostelium developme
A;Reference number: S71629; MUID:96324397; PMID:8670894
A;Accession: S71629
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-2150 <NAN>
A;Cross-references: UNIPROT:Q23863; EMBL:U42597
A;Experimental source: strain Ax4
C;Genetics:
A;Gene: dhka
A;Map position: 6
C;Keywords: autophosphorylation; phosphoprotein; phosphotransferase; two-component regul
F;2076/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 51.9%; Score 42; DB 2; Length 2150;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPQNC 15
|||::|::|
Db 345 GNNSPRSCGSGNC 357

RESULT 9
AH2488
hypothetical protein alr7088 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120al
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2488
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2488
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-148 <KUR>
A;Cross-references: UNIPROT:Q8YL47; GB:BA000020; PIDN:BA078172.1; PID:g17135626; GSPDB:G
A;Experimental source: strain PCC 7120
```

```
C;Genetics:
A;Gene: alr7088
A;Genome: plasmid

Query Match 50.6%; Score 41; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TQSPQNC 15
|||||
Db 85 TQSPQNC 91

RESULT 10
AI2098
hypothetical protein all2344 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AI2098
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AI2098
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-439 <KUR>
A;Cross-references: UNIPROT:Q8YUK2; GB:BA000019; PIDN:BA074043.1; PID:g17131436; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2344

Query Match 50.6%; Score 41; DB 2; Length 439;
Best Local Similarity 70.0%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSP 12
|||||::|
Db 257 GNSSPTTENP 266

RESULT 11
T02074
4-coumarate-CoA ligase (EC 6.2.1.12) - common tobacco
N;Alternate names: 4-coumaroyl-CoA synthetase
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: T02074
R;Katayama, Y.; Kawai, S.; Morohoshi, N.; Kajita, S.
submitted to the EMBL Data Library, December 1994
A;Description: Cloning and nucleotide sequence of 4-coumarate:coenzyme A ligase gene from
A;Reference number: Z14545
A;Accession: T02074
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-542 <NAT>
A;Cross-references: UNIPROT:Q42943; EMBL:DA3773; PIDN:BA07828.1
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology
C;Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
F;73-532/Domain: acetate-CoA ligase homology <ACL>

Query Match 50.6%; Score 41; DB 1; Length 542;
Best Local Similarity 61.5%; Pred. No. 52;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPQNC 15
|||||::|
Db 377 GNSLPNQSGEIC 389

RESULT 12
```

T03789

4-coumarate-CoA ligase (EC 6.2.1.12) 4CL2 - common tobacco  
 N:Alternate names: 4-coumaroyl-CoA synthetase  
 C:Species: Nicotiana tabacum (common tobacco)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C:Accession: T03789  
 R:Lee, D.; Douglas, C.J.  
 Plant Physiol. 112, 193-205, 1996  
 A:Title: Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene family  
 A:Reference number: Z15086; MUID:96416441; PMID:8819324  
 A:Accession: T03789  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-542 <LEE>  
 A:Cross-references: UNIPROT:O24146; EMBL:U50846; NID:g1663723; PIDN:AAB18638.1; PID:g1663723

C:Genetics:

A:Gene: 4CL2

C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology  
 C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis  
 P,7,3-532/Domain: acetate-CoA ligase homology <ACL>

Query Match 50.6%; Score 41; DB 2; Length 542;  
 Best Local Similarity 61.5%; Pred. No. 52;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GNSSPRTSQPNC 15

Db 377 GNSLPRNQSGETC 389

RESULT 13

153395  
 antigen LEC-A - mouse  
 C:Species: Mus sp. (mouse)  
 C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
 C:Accession: I53395  
 R:de Bergcyck, V.; De Plaen, E.; Chomez, P.; Boon, T.; Van Pel, A.  
 Eur. J. Immunol. 24, 2203-2212, 1994  
 A:Title: An intracisternal A-particle sequence codes for an antigen recognized by syngeneic cells  
 A:Reference number: I53395; MUID:94374432; PMID:8088336  
 A:Accession: I53395  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-599 <RES>  
 A:Cross-references: UNIPROT:Q64033; GB:S74315; NID:g765347; PIDN:AAB31958.1; PID:g765348

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

Query Match 50.6%; Score 41; DB 2; Length 599;  
 Best Local Similarity 63.6%; Pred. No. 57;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GNSSPRTSQPQ 13

Db 556 GSSGPRSQGPQ 566

RESULT 14

1531034  
 retrovirus-related gag polyprotein - mouse intracisternal A-particle M1AB12 (fragment)  
 C:Species: mouse intracisternal A-particle M1AB12  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-Feb-1997  
 C:Accession: S31034  
 R:Grossman, Z.; Mietz, J.A.; Kuff, E.L.  
 Nucleic Acids Res. 15, 3823-3834, 1987  
 A:Title: Nearly identical members of the heterogeneous IAP gene family are expressed in mouse intracisternal A-particle virus  
 A:Reference number: S31034; MUID:87231071; PMID:2884637  
 A:Accession: S31034  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-717 <GRO>  
 A:Cross-references: EMBL:X05545

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: glycoprotein; polyprotein

P,7,2,142,339,413,555/Binding site: carbohydrate (Asn) (covalent) #status predicted

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: glycoprotein; polyprotein  
 P,7,2,142,339,413,555/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.6%; Score 41; DB 2; Length 717;

Best Local Similarity 63.6%; Pred. No. 69;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GNSSPRTSQPQ 13

Db 414 GSSGPRSQGPQ 424

RESULT 15

1531035  
 retrovirus-related gag polyprotein - mouse intracisternal A-particle M1AD8 (fragment)  
 C:Species: mouse intracisternal A-particle M1AD8  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-Feb-1997  
 C:Accession: S31035  
 R:Grossman, Z.; Mietz, J.A.; Kuff, E.L.  
 Nucleic Acids Res. 15, 3823-3834, 1987  
 A:Title: Nearly identical members of the heterogeneous IAP gene family are expressed in mouse intracisternal A-particle virus  
 A:Reference number: S31034; MUID:87231071; PMID:2884637  
 A:Accession: S31035  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-717 <GRO>  
 A:Cross-references: EMBL:X05546

C:Genetics:

A:Gene: gag

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: glycoprotein; polyprotein

P,7,2,142,339,413,555/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.6%; Score 41; DB 2; Length 717;

Best Local Similarity 63.6%; Pred. No. 69;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GNSSPRTSQPQ 13

Db 414 GSSGPRSQGPQ 424

RESULT 16

1543742  
 female sterile homeotic protein, 205K - fruit fly (Drosophila melanogaster)  
 N:Alternate names: membrane protein fish, 205K  
 N:Contains: female sterile homeotic protein, 110K  
 C:Species: Drosophila melanogaster  
 C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 09-Jul-2004  
 C:Accession: A43742; B43742  
 R:Haynes, S.R.; Mozer, B.A.; Bhatia-Dey, N.; Dawid, I.B.  
 Dev. Biol. 134, 246-257, 1989  
 A:Title: The Drosophila fish locus, a maternal effect homeotic gene, encodes apparent men  
 A:Reference number: A43742; MUID:89276730; PMID:2567251  
 A:Accession: A43742  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2038 <HAY>  
 A:Cross-references: UNIPROT:P13709; EMBL:M23221; NID:g157452; PIDN:AAA28540.1; PID:g1574

C:Genetics:

A:Gene: fish

A:Cross-references: FlyBase:FBgn0004656

C:Superfamily: unassigned bromodomain proteins; bromodomain homology

C:Keywords: alternative splicing; transmembrane protein

P,1-2038/Product: female sterile homeotic protein, 205K #status predicted &lt;MA2&gt;

F:1-1106/Product: female sterile homeotic protein, 110K #status predicted <MAT>  
F:59-1116/Domain: bromodomain homology <BRO1>  
F:503-560/Domain: bromodomain homology <BRO2>

Query Match 50.6%; Score 41; DB 2; Length 2038;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPQN 14

|||||

Db 1985 QQSPPAQQSPQD 1996

RESULT 17

S71466 homeotic protein SOX-12 protein - African clawed frog

N/Alternate names: SRY-related protein

C/Species: Xenopus laevis (African clawed frog)

C/Date: 09-Dec-1997 #sequence\_revision 09-Dec-1997 #text\_change 16-Aug-2004

C/Accession: S71466; S22947; S21490

F/Komatsu, N.; Hiraoka, Y.; Shiozawa, M.; Ogawa, M.; Aiso, S.

Biochim. Biophys. Acta 1305, 117-119, 1996

A/Title: Cloning and expression of Xenopus laevis xSox12 cDNA.

A/Reference number: S71466; MUID:96180312; PMID:8597594

A/Accession: S71466

A/Molecule type: mRNA

A/Residues: 1-470 <KOM>

A/Cross-references: UNIPROT:P40649; EMBL:D50552; NID:g1255903; PIDN:BA09119.1; PID:d100

A/Experimental source: ovary

R/Denny, P.; Swift, S.; Brand, N.; Dabhadre, N.; Barton, P.; Ashworth, A.

Nucleic Acids Res. 20, 2887, 1992

A/Title: A conserved family of genes related to the testis determining gene, SRY.

A/Reference number: S22947; MUID:92310993; PMID:1614875

A/Accession: S22947

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 308-359 <DEN>

A/Cross-references: EMBL:X65655; NID:g65261; PIDN:CAA46606.1; PID:g938305

C/Genetics:

A/Gene: sox12

C/Superfamily: HMG box homology

C/Keywords: DNA binding; leucine zipper

F:35-54/Region: leucine zipper

F:294-367/Domain: HMG box homology <HMG1>

Query Match 50.0%; Score 40.5; DB 2; Length 470;  
Best Local Similarity 66.7%; Pred. No. 55;  
Matches 10; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Qy 1 LLGNSSPRTQSPQNC 15

|||||

Db 415 LLGNLS---QSPQEC 426

RESULT 18

E82515

hypothetical protein XF2717 [imported] - Xylella fastidiosa (strain 9a5c)

C/Species: Xylella fastidiosa

C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C/Accession: E82515

R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A/Reference number: A82515; MUID:20365717; PMID:10910347

A/Note: for a complete list of authors see reference number A59328 below

A/Accession: E82515

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-94 <SIM>

A/Cross-references: UNIPROT:Q9PA03; GB:AE004078; GB:AE003849; NID:g9187832; PIDN:AAF8575

A/Experimental source: strain 9a5c

R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H

as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A/Reference number: A59328  
A/Contents: annotation  
C/Genetics:  
A/Gene: XP2717

Query Match 49.4%; Score 40; DB 2; Length 94;  
Best Local Similarity 42.9%; Pred. No. 13;  
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQN 14

|||||

Db 5 VIGNHTPTTSPES 18

RESULT 19

S56155

sulfocyanin - Sulfolobus acidocaldarius

N/Alternate names: blue copper protein

C/Species: Sulfolobus acidocaldarius

C/Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004

C/Accession: S56155

R/Castresana, J.; Luebben, M.; Saraste, M.

J. Mol. Biol. 250, 202-210, 1995

A/Title: New archaeobacterial genes coding for redox proteins: implications for the evolu

A/Reference number: S56155; MUID:95333177; PMID:7608970

A/Accession: S56155

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-199 <CAS>

A/Cross-references: UNIPROT:Q53765; EMBL:Z48338; NID:g927521; PID:g927523

F:7-34/Domain: transmembrane #status predicted <TRM>

F:110,171,176,181/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 49.4%; Score 40; DB 2; Length 199;  
Best Local Similarity 61.5%; Pred. No. 28;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQ 13

|||||

Db 114 LLONSTPTQSPPE 126

RESULT 20

H90369

2-haloalkanoic acid dehalogenase [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C/Accession: H90369

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P

arrest, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: H90369

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-212 <KUR>

A/Cross-references: UNIPROT:Q97WTG; GB:AE006641; NID:g13815309; PIDN:AAK42215.1; GSPDB:G

C/Genetics:

A/Gene: SSO2028

Query Match

49.4%; Score 40; DB 2; Length 212;

```

Best Local Similarity 80.0%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQ 10
DB 110 LLGNSSPRTK 119

RESULT 21
I51916
H+/K+-exchanging ATPase (EC 3.6.3.10) alpha chain - rat (fragment)
C:Species: Rattus sp. (rat)
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 19-Apr-2002
A:Accession: I51916
R:Ahn, K.Y.; Kone, B.C.
Am. J. Physiol. 268, F99-F109, 1995
A:Title: Expression and cellular localization of mRNA encoding the 'gastric' isoform of
A:Reference number: I51916; MUID:95142238; PMID:7840253
A:Accession: I51916
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-454 <RES>
A:CROSS-references: GB:S74801; NID:9802079
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase; phosphoprotein
F:332/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 49.4%; Score 40; DB 2; Length 454;
Best Local Similarity 53.8%; Pred. No. 64;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQ 13
DB 174 LTGESEPTQTSPE 186

RESULT 22
A12964
Hypothetical protein Atu322 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A:Accession: A12964
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: A12964
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-551 <KUR>
A:CROSS-references: UNIPROT:Q8UAQ1; GB:AE008689; PIDN:AAL44135.1; PID:g17741707; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3322
A:Map position: linear chromosome

Query Match 49.4%; Score 40; DB 2; Length 551;
Best Local Similarity 61.5%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQ 13
DB 536 LSGNGRPAQSPQ 548

RESULT 23
ABC protein AGR_L_3000 (AB015053) [imported] - Agrobacterium tumefaciens (strain C58, Ce
C:Species: Agrobacterium tumefaciens

```

```

C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: D98318
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D98318
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-590 <KUR>
A:CROSS-references: UNIPROT:Q8UAQ1; GB:AE007870; PIDN:AAK90070.1; PID:g15160053; GSPDB:B
C:Genetics:
A:Gene: AGR_L_3000
A:Map position: linear chromosome

Query Match 49.4%; Score 40; DB 2; Length 690;
Best Local Similarity 61.5%; Pred. No. 97;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQ 13
DB 675 LSGNGRPAQSPQ 687

RESULT 24
A37227
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - chicken
C:Species: Gallus gallus (chicken)
C>Date: 16-Sep-1992 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
A:Accession: I50394; A37227
R:Takeyasu, K.; Lemas, M.; Fambrough, D.M.
Am. J. Physiol. 259, 619-630, 1991
A:Title: Stability of the Na+, K+-ATPase alpha-subunit isoforms in evolution.
A:Reference number: I50394
A:Accession: I50394
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1017 <TA2>
A:CROSS-references: UNIPROT:P24797; GB:M59599; NID:g212405; PIDN:AAA48981.1; PID:g212406
R:Takeyasu, K.; Lemas, V.; Fambrough, D.M.
Am. J. Physiol. 259, C619-C630, 1990
A:Title: Stability of Na(+)-K(+)-ATPase alpha-subunit isoforms in evolution.
A:Reference number: A37227; MUID:91023019; PMID:2171348
A:Accession: A37227
A:Molecule type: mRNA
A:Residues: 3-1017 <TA2>
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; glycoprotein; hydrolase; phosphoprotein
F:581-777/Domain: ATPase nucleotide-binding domain homology <ATN>
F:210,478/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:371/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 49.4%; Score 40; DB 2; Length 1017;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLGNSSPRTQSPQ 13
DB 213 LTGESEPTQTSPE 225

RESULT 25
A34474
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - human
N:Alternate names: Na+/K+-exchanging ATPase alpha chain-4; sodium/potassium transporting
C:Species: Homo sapiens (man)
C>Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 09-Jul-2004
A:Accession: A34474; B27795; D27397
R:Shull, M.M.; Pugh, D.G.; Lingrel, J.B.
J. Biol. Chem. 264, 17532-17543, 1989
A:Title: Characterization of the human Na,K-ATPase alpha2 gene and identification of int
A:Reference number: A34474; MUID:90008924; PMID:2477373

```

A;Accession: A34474  
A;Molecule type: DNA  
A;Residues: 1-1020 <SHU>  
A;Cross-references: UNIPROT:P50993; GB:J05096; NID:g179164; PIDN:AAA51797.1; PID:g179165  
R;Shull, M.M.; Lingrel, J.B.  
Proc. Natl. Acad. Sci. U.S.A. 84, 4039-4043, 1987  
A;Title: Multiple genes encode the human Na<sup>+</sup>/K<sup>+</sup>-ATPase catalytic subunit.  
A;Reference number: A94158; MUID:872331946; PMID:3035563  
A;Accession: B27795  
A;Molecule type: DNA  
A;Residues: 211-249 <SH2>  
A;Cross-references: GB:M16795; NID:g179196; PIDN:AAA51799.1; PID:g553194  
R;Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkaryov, Y.A.; Allikmets, R.L.; M  
tina, M.B.; Sverdlov, V.B.; Modyanov, N.N.; Ovchinnikov, Y.A.  
FEBS Lett. 217, 275-278, 1987  
A;Title: The family of human Na<sup>+</sup>/K<sup>+</sup>-ATPase genes. No less than five genes and/or pseudoge  
A;Reference number: A27397; MUID:87247232; PMID:3036582  
A;Accession: D27397  
A;Molecule type: DNA  
A;Residues: 251-442 <SVE>  
A;Cross-references: GB:M27571  
C;Genetics:  
A;Gene: GDB:ATPIA2  
A;Cross-references: GDB:119712; OMIM:182340  
A;Map position: 1q21-1q23  
C;Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp  
F;6-1020/Product: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha-2 chain #status predicted <MAT>  
F;6-93/Domain: intracellular #status predicted <INT1>  
F;128-147/Domain: transmembrane #status predicted <TM1>  
F;148-288/Domain: intracellular #status predicted <INT2>  
F;289-311/Domain: transmembrane #status predicted <TM2>  
F;318-346/Domain: intracellular #status predicted <INT3>  
F;347-783/Domain: transmembrane #status predicted <TM3>  
F;584-780/Domain: ATPase nucleotide-binding domain homology <ATN>  
F;784-807/Domain: transmembrane #status predicted <TM5>  
F;846-871/Domain: transmembrane #status predicted <TM6>  
F;872-949/Domain: intracellular #status predicted <INT4>  
F;950-975/Domain: transmembrane #status predicted <TM7>  
F;976-1020/Domain: extracellular #status predicted <EXT>  
F;374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F;505/Binding site: ATP (Lys) #status predicted  
F;714,718,723/Active site: Asp, Asp, Lys #status predicted  
P;714,718,723/Active site: Asp, Asp, Lys #status predicted  
Query Match 49.4%; Score 40; DB 2; Length 1020;  
Best Local Similarity 53.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 LLGNSSPRTSQPQ 13  
Db 216 LTGESEPTQTSPE 228  
RESULT 26  
B24639  
Na<sup>+</sup>/K<sup>+</sup>-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - rat  
N;Alternate names: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha-plus chain  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
C;Accession: B24639  
R;Shull, G.B.; Greab, J.; Lingrel, J.B.  
Biochemistry 25, 8125-8132, 1986  
A;Title: Molecular cloning of three distinct forms of the Na<sup>+</sup>/K<sup>+</sup>-ATPase alpha-subunit fr  
A;Reference number: A30512; MUID:87128908; PMID:3028470  
A;Accession: B24639  
A;Molecule type: mRNA  
A;Residues: 1-1020 <SHU>  
A;Cross-references: UNIPROT:P06686; EMBL:M14512; NID:g203028; PIDN:AAA40776.1; PID:g2030  
C;Genetics:  
A;Gene: NKAA2  
C;Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp

F;6-1020/Product: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha-2 chain #status predicted <MAT>  
F;6-93/Domain: intracellular #status predicted <INT1>  
F;94-119/Domain: transmembrane #status predicted <TM1>  
F;128-147/Domain: transmembrane #status predicted <TM2>  
F;148-288/Domain: intracellular #status predicted <INT2>  
F;289-311/Domain: transmembrane #status predicted <TM3>  
F;318-346/Domain: transmembrane #status predicted <TM4>  
F;347-783/Domain: intracellular #status predicted <INT3>  
F;584-780/Domain: ATPase nucleotide-binding domain homology <ATN>  
F;784-807/Domain: transmembrane #status predicted <TM5>  
F;846-871/Domain: transmembrane #status predicted <TM6>  
F;872-949/Domain: intracellular #status predicted <INT4>  
F;950-975/Domain: transmembrane #status predicted <TM7>  
F;976-1020/Domain: extracellular #status predicted <EXT>  
F;374/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F;505/Binding site: ATP (Lys) #status predicted  
F;714,718,723/Active site: Asp, Asp, Lys #status predicted  
Query Match 49.4%; Score 40; DB 2; Length 1020;  
Best Local Similarity 53.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 LLGNSSPRTSQPQ 13  
Db 216 LTGESEPTQTSPE 228  
RESULT 27  
I51674  
gastric H(+)-K(+)-ATPase alpha-subunit - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C;Accession: I51674  
R;Mathews, P.M.; Claeys, D.; Jaissner, F.; Geering, K.; Kraehenbuhl, J.D.; Kraehenbuhl, J  
Am. J. Physiol. 268, C1207-C1214, 1995  
A;Title: Primary structure and functional expression of the mouse and frog alpha-subunit  
A;Reference number: I49143; MUID:95282844; PMID:7762614  
A;Accession: I51674  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-1031 <MAT>  
A;Cross-references: UNIPROT:Q92126; EMBL:U17249; NID:g596055; PIDN:AAA76601.1; PID:g5960  
C;Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain i  
C;Keywords: ATP; phosphoprotein  
F;584-790/Domain: ATPase nucleotide-binding domain homology <ATN>  
F;383/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F;515/Binding site: ATP (Lys) #status predicted  
Query Match 49.4%; Score 40; DB 2; Length 1031;  
Best Local Similarity 53.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 LLGNSSPRTSQPQ 13  
Db 225 LTGESEPTQTSPE 237  
RESULT 28  
I49143  
gastric H(+)-K(+)-ATPase alpha subunit - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I49143  
R;Mathews, P.M.; Claeys, D.; Jaissner, F.; Geering, K.; Kraehenbuhl, J.D.; Kraehenbuhl, J  
Am. J. Physiol. 268, C1207-C1214, 1995  
A;Title: Primary structure and functional expression of the mouse and frog alpha-subunit  
A;Reference number: I49143; MUID:95282844; PMID:7762614  
A;Accession: I49143  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-1033 <RES>  
A;Cross-references: UNIPROT:Q64436; EMBL:U17282; NID:g596067; PIDN:AAA79514.1; PID:g5960  
C;Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain ;

A>Title: Molecular cloning and structural analysis of canine gastric H<sup>+</sup>, K<sup>+</sup>-ATPase.  
A:Reference number: JN0903; MUID:94071880; PMID:8250881  
A:Accession: JN0903  
A:Molecule type: mRNA  
A:Residues: 1-1034 <SON>  
A:Experimental source: stomach  
A>Note: The authors translated the codon GGG for residue 728 as nothing  
A>Note: the DNA sequence is also determined and has 21 introns  
C:Comment: This enzyme functions as a proton pump in the secretion of hydrochloric acid  
C:Genetics:  
A:Introns: 4/3; 51/3; 71/3; 139/3; 177/3; 262/1; 351/3; 418/1; 454/3; 499/3; 564/1; 623/  
C:Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; glycoprotein; hydrolase; ion transport; phosphoprotein; potassium trans;  
F:597-793/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:225,493/Binding site: carbohydurate (Asn) (covalent) #status predicted  
F:386/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:518/Binding site: ATP (Lys) #status predicted

Query Match 49.4%; Score 40; DB 2; Length 1034;  
Best Local Similarity 53.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRQTSPQ 13  
| | | : | | |  
Db 228 LTGESEPTQRSPE 240

RESULT 31  
A31671  
H+/K+-exchanging ATPase (EC 3.6.3.10) alpha chain - pig  
N/Alternate names: H+/K+-transporting ATPase  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 21-May-1990 #sequence revision 21-May-1990 #text\_change 09-Jul-2004  
C/Accession: A31671; A24228; S21269; S18375  
R/Maeda, M.; Ishizaki, J.; Futai, M.  
Biochem. Biophys. Res. Commun. 157, 203-209, 1988  
A>Title: cDNA cloning and sequence determination of pig gastric (H<sup>+</sup>) + K(+) )-ATPase.  
A:Reference number: A31671; MUID:89061712; PMID:2848518  
A:Accession: A31671  
A:Molecule type: mRNA  
A:Residues: 1-1034 <NAE>  
A:Cross-references: UNIPROT:P19156; GB:M22724; NID:g164383; PIDN:AAA31003.1; PID:g164384  
R/Lane, L.K.; Kirley, T.L.; Ball Jr., W.J  
Biochem. Biophys. Res. Commun. 138, 185-192, 1986  
A:Reference number: A24228; MUID:86295667; PMID:3017315  
A:Accession: A24228  
A:Molecule type: protein  
A:Residues: 'Y', 3-18 <LAN>  
R/Callaghan, J.M.; Tob, B.H.; Simpson, R.J.; Baldwin, G.S.; Gleeson, P.A.  
Biochem. J. 283, 63-68, 1992  
A>Title: Rapid purification of the gastric H(+)/K(+)-ATPase complex by tomato-lectin aff  
A:Reference number: S21269; MUID:92231849; PMID:1314570  
A:Accession: S21269  
A:Molecule type: protein  
A:Residues: 173-182;777-783;1022-1033 <CAL>  
R/van Uem, T.J.F.; Swarts, H.G.P.; de Pont, J.J.H.H.M.  
Biochem. J. 280, 243-248, 1991  
A>Title: Determination of the epitope for the inhibitory monoclonal antibody 5-B6 on the  
A:Reference number: S18375; MUID:92074982; PMID:1720615  
A:Accession: S18375  
A:Molecule type: protein  
A:Residues: 'X', 49-53, 'X', 55-58, 'X', 380, 'X', 382-383, 'XX', 386, 'X'; 457-458, 'X', 460-461 <VA  
C:Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; potassium transport; transmem  
F:597-793/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:225,493/Binding site: carbohydurate (Asn) (covalent) #status predicted  
F:386/Active site: Asp (aspartylphosphate intermediate) #status experimental  
F:518/Binding site: ATP (Lys) #status predicted

Query Match 49.4%; Score 40; DB 2; Length 1034;  
Best Local Similarity 53.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1  LLGNSSPRTQSPQ 13
Db      228  LTGESEPTQTRSPE 240

RESULT 32
A35292
H+/K+-exchanging ATPase (EC 3.6.3.10) alpha chain - human
N;Alternate names: H+/K+-transporting ATPase alpha chain
C;Species: Homo sapiens (man)
C;Date: 17-Aug-1990 #sequence_revision 06-Nov-1992 #text_change 09-Jul-2004
C;Accession: A36558; A35292
R;Newman, P.R.; Greeb, J.; Keeton, T.P.; Reyes, A.A.; Shull, G.E.
DNA Cell Biol. 9, 749-762, 1990
A;Title: Structure of the human gastric H,K-ATPase gene and comparison of the 5'-flankin
A;Reference number: A36558; MUID:91090845; PMID:2176086
A;Accession: A36558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1035 <NEW>
A;Cross-references: UNIPROT:P20648; GB:M63962; NID:g184106; PIDN:AAA35988.1; PID:g184107
F;Maeda, M.; Oshiman, K.I.; Tamura, S.; Futai, M.
J. Biol. Chem. 265, 9027-9032, 1990
A;Title: Human Gastric H(+) + K(+) -ATPase gene. Similarity to (Na(+) + K(+)) -ATPase gen
A;Reference number: A35292; MUID:90264383; PMID:2160952
A;Accession: A35292
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1035 <MAE>
A;Cross-references: GB:J05451; NID:g561633; PIDN:AAA51010.1; PID:g561634; GB:J05452
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; transmembrane protein
F;598-794/Domain: ATPase nucleotide-binding domain homology <ATN>
F;226,494/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;387/Active site: Asp (aspartylphosphate intermediate) #status predicted
F;519/Binding site: ATP (Lys) #status predicted

Query Match      49.4%; Score 40; DB 2; Length 1035;
Best Local Similarity 53.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1  LLGNSSPRTQSPQ 13
Db      229  LTGESEPTQTRSPE 241

RESULT 33
S23406
H+/K+-exchanging ATPase (EC 3.6.3.10) alpha chain - rabbit
N;Alternate names: H+/K+-transporting ATPase alpha chain
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C;Accession: S23406
R;Bamberg, K.; Mercier, F.; Reuben, M.A.; Kobayashi, Y.; Munson, K.B.; Sachs, G.
Biochim. Biophys. Acta 1131, 69-77, 1992
A;Title: cDNA cloning and membrane topology of the rabbit gastric H(+)/K(+)-ATPase alpha
A;Reference number: S23406; MUID:92256493; PMID:1316171
A;Accession: S23406
A;Molecule type: mRNA
A;Residues: 1-1035 <BAM>
A;Cross-references: UNIPROT:P27112; EMBL:X64694; NID:g1470; PIDN:CAA45927.1; PID:g1471
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; transmembrane protein
F;598-794/Domain: ATPase nucleotide-binding domain homology <ATN>
F;226,494/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;387/Active site: Asp (aspartylphosphate intermediate) #status predicted
F;519/Binding site: ATP (Lys) #status predicted

Query Match      49.4%; Score 40; DB 2; Length 1035;
Best Local Similarity 53.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1  LLGNSSPRTQSPQ 13
```

```
Db      229  LTGESEPTQTRSPE 241

RESULT 34
S20484
hypothetical protein 2 - Azospirillum brasilense
C;Species: Azospirillum brasilense
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S20484
R;Vieille, C.; Elmerich, C.
Mol. Gen. Genet. 231, 375-384, 1992
A;Title: Characterization of an Azospirillum brasilense Sp7 gene homologous to Alcaligen
A;Reference number: S20483; MUID:92167956; PMID:1538694
A;Accession: S20484
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <VIE>
A;Cross-references: UNIPROT:Q00644; EMBL:X64772; NID:g38684; PIDN:CAA46020.1; PID:g58069;
C;Genetic8;
A;Start codon: GTG

Query Match      48.1%; Score 39; DB 2; Length 169;
Best Local Similarity 57.1%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      2  LLGNSSPRTQSPQNC 15
Db      6  LFNSSPRAGGPSAC 19

RESULT 35
A45056
osteogenic protein 2 precursor - human
N;Alternate names: bone morphogenetic protein 8; OP-2
C;Species: Homo sapiens (man)
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45056
R;Ozskynak, E.; Schnegelsberg, P.N.; Jin, D.F.; Clifford, G.M.; Warren, F.D.; Drier, E.J.
J. Biol. Chem. 267, 25220-25227, 1992
A;Title: Osteogenic protein-2. A new member of the transforming growth factor-beta super:
A;Reference number: A45056; MUID:93094231; PMID:1460021
A;Accession: A45056
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-402 <OZK>
A;Cross-references: UNIPROT:P34820; GB:M97016; NID:g189389; PIDN:AAB01360.1; PID:g189390
A;Experimental source: hippocampus
A;Note: sequence extracted from NCBI backbone (NCBIP:120189)
C;Genetics;
A;Gene: GDB:EMP8; OP-2
A;Cross-references: GDB:I36392
C;Superfamily: inhibin

Query Match      48.1%; Score 39; DB 2; Length 402;
Best Local Similarity 58.3%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1  LLGNSSPRTQSP 12
Db      235  LLGRAPRSQOP 246

RESULT 36
B46165
envelope surface glycoprotein SU - feline leukemia virus (strain FA27)
C;Species: feline leukemia virus
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: B46165
R;Brojatsch, J.; Kristal, B.S.; Viglianti, G.A.; Khiroya, R.; Hoover, E.A.; Mullins, J.I.
Proc. Natl. Acad. Sci. U.S.A. 89, 8457-8461, 1992
A;Title: Feline leukemia virus subgroup C phenotype evolves through distinct alterations
A;Reference number: A46165; MUID:92409535; PMID:1326757
```



A;Title: Nucleotide sequence analysis of the LTRs and env genes of SM-FesV and GA-FesV.  
A;Reference number: A33741; MUID:88044502; PMID:2823466  
A;Accession: B33741  
A:Molecule type: DNA  
A;Residues: 1-642 <GUI>  
C;Genetics:  
A:Gene: env  
C;Superfamily: type C retrovirus env polyprotein  
C;Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protein  
F;1-32/Domain: signal sequence #status predicted <SIG>  
E;33-445/Product: coat protein gp70 #status predicted <GUP>  
F;446-642/Product: coat protein p15E #status predicted <GUQ>  
F;34, 42, 57, 267, 302, 307, 331, 334, 374, 390, 410, 578/Binding site: carbohydrate (Asn) (covalent)

Query Match 48.1%; Score 39; DB 1; Length 642;  
Best Local Similarity 46.7%; Pred. No. 1.3e+02;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
:| | | | :| :| :|  
Db 328 ILGNYSNQTNPSPC 342

RESULT 42  
T10533  
env polyprotein precursor - feline leukemia virus (strain FeLV-FAIDS)  
N;Contains: env protein gp70; env protein p15E  
C;Species: feline leukemia virus  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C;Accession: T10533  
J;Donahue, P.R.; Hoover, E.A.; Beltz, G.A.; Riedel, N.; Hirsch, V.M.; Overbaugh, J.; Mulla, R.;  
J. Virol. 62, 722-731, 1988  
A;Title: Strong sequence conservation among horizontally transmissible, minimally pathogenic  
A;Reference number: Z17078; MUID:88119207; PMID:2828667  
A;Accession: T10533  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A;Residues: 1-642 <DON>  
A;Cross-references: UNIPROT:O85522; EMBL:M18247; NID:g3233904; PIDN:AAA93093.1; PID:g3233904  
C;Superfamily: type C retrovirus env polyprotein

Query Match 48.1%; Score 39; DB 2; Length 642;  
Best Local Similarity 46.7%; Pred. No. 1.3e+02;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
:| | | | :| :| :|  
Db 328 ILGNYSNQTNPSPC 342

RESULT 43  
VCMVSS  
env polyprotein precursor - feline sarcoma virus (strain SM)  
N;Contains: coat protein gp70; coat protein p15E  
C;Species: feline sarcoma virus  
A;Note: host Felis silvestris catus (domestic cat)  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 12-Apr-1996  
C;Accession: A33741  
R;Guilhot, S.; Hampe, A.; D'Auriol, L.; Galibert, F.  
Virology 161, 252-258, 1987  
A;Title: Nucleotide sequence analysis of the LTRs and env genes of SM-FesV and GA-FesV.  
A;Reference number: A33741; MUID:88044502; PMID:2823466  
A;Accession: A33741  
A:Molecule type: DNA  
A;Residues: 1-645 <GUI>  
C;Genetics:  
A:Gene: env  
C;Superfamily: type C retrovirus env polyprotein  
C;Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protein  
F;1-36/Domain: signal sequence #status predicted <SIG>  
F;37-448/Product: coat protein gp70 #status predicted <GUP>  
F;449-645/Product: coat protein p15E #status predicted <GUQ>  
F;38, 46, 61, 270, 305, 310, 334, 337, 377, 393, 413, 591/Binding site: carbohydrate (Asn) (covalent)



Db 842 IGNSNHGSQSPRN 854

Search completed: August 25, 2005, 11:44:35  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 11:38:14 ; Search time 161 Seconds  
(without alignments)  
36.608 Million cell updates/sec

Title: US-10-803-541-2

Perfect score: 81

Sequence: 1 LLGNSPRQSPQNC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	81	100.0	15	US-10-803-541-2
2	81	100.0	17	US-10-803-541-20
3	81	100.0	515	US-09-315-355-46
4	81	100.0	515	US-10-848-572-46
5	81	100.0	664	US-10-408-765A-2079
6	81	100.0	664	US-10-473-127-1658
7	81	100.0	664	US-10-943-400-2
8	81	100.0	664	US-10-803-541-4
9	69	85.2	15	US-10-803-541-16
10	69	85.2	18	US-10-803-541-21
11	69	85.2	665	US-10-803-541-9

Sequence 132511,	16	US-10-437-963-132511	162	55.6	45	12
Sequence 198650,	16	US-10-425-115-198650	122	54.3	44	13
Sequence 145179,	15	US-10-424-599-145179	275	54.3	44	14
Sequence 2862, Ap	15	US-10-094-749-2862	286	54.3	44	15
Sequence 117902,	16	US-10-437-963-117902	826	54.3	44	16
Sequence 202808,	17	US-10-437-963-202808	539	53.1	43	17
Sequence 392, App	18	US-10-481-032A-392	558	53.1	43	18
Sequence 203613,	16	US-10-437-963-203613	1797	53.1	43	19
Sequence 203615,	20	US-10-437-963-203615	531	53.1	43	20
Sequence 46748, A	20	US-10-767-701-46748	202	52.5	42	21
Sequence 234564,	81	US-10-424-599-234564	51.9	51.9	42	22
Sequence 324347,	16	US-10-425-115-324347	51.9	51.9	42	23
Sequence 20, Appl	15	US-10-283-940-20	175	51.9	42	24
Sequence 19, Appl	25	US-11-022-454-20	175	51.9	42	25
Sequence 19, Appl	26	US-09-994-064-19	177	51.9	42	26
Sequence 19, Appl	27	US-09-993-777-19	177	51.9	42	27
Sequence 206208,	28	US-10-836-383-19	177	51.9	42	28
Sequence 3255, Ap	29	US-10-424-599-206208	183	51.9	42	29
Sequence 203609,	30	US-10-104-047-3255	256	51.9	42	30
Sequence 204611,	31	US-10-437-963-203609	1701	51.9	42	31
Sequence 203618,	32	US-10-437-963-204611	1828	51.9	42	32
Sequence 17, Appl	33	US-10-437-963-203618	1980	51.9	42	33
Sequence 218829,	34	US-10-135-322-17	2150	51.9	42	34
Sequence 352933,	35	US-10-425-115-218829	78	50.6	41	35
Sequence 230506,	36	US-10-425-115-230506	86	50.6	41	36
Sequence 57287, A	37	US-10-424-599-230506	112	50.6	41	37
Sequence 38640, A	38	US-10-767-701-57287	112	50.6	41	38
Sequence 47848, A	39	US-10-425-114-38640	243	50.6	41	39
Sequence 60855, A	40	US-10-425-114-47848	310	50.6	41	40
Sequence 69314, A	41	US-10-425-114-69314	316	50.6	41	41
Sequence 257256,	42	US-10-425-115-257256	320	50.6	41	42
Sequence 258656,	43	US-10-425-115-258656	334	50.6	41	43
Sequence 258661,	44	US-10-425-115-258661	334	50.6	41	44
Sequence 257255,	45	US-10-425-115-257255	343	50.6	41	45
Sequence 58835, A	46	US-10-425-114-58835	354	50.6	41	46
Sequence 238659, A	47	US-10-425-115-238659	364	50.6	41	47
Sequence 60636, A	48	US-10-425-114-60636	365	50.6	41	48
Sequence 59471, A	49	US-10-425-114-59471	370	50.6	41	49
Sequence 67675, A	50	US-10-425-114-67675	370	50.6	41	50
Sequence 67309, A	51	US-10-425-114-67309	378	50.6	41	51
Sequence 189698,	52	US-10-424-599-189698	100	50.0	40	52
Sequence 1261, Ap	53	US-10-862-195-1261	10	49.4	40	53
Sequence 7568, Ap	54	US-10-156-761-7568	54	49.4	40	54
Sequence 351554,	55	US-10-425-115-351554	54	49.4	40	55
Sequence 297429,	56	US-10-425-115-297429	81	49.4	40	56
Sequence 269275,	57	US-10-424-599-269275	104	49.4	40	57
Sequence 284701,	58	US-10-425-115-284701	120	49.4	40	58
Sequence 4641, Ap	59	US-10-108-260A-4641	123	49.4	40	59
Sequence 54437, A	60	US-10-425-115-233052	170	49.4	40	60
Sequence 233052,	61	US-10-767-701-54437	197	49.4	40	61
Sequence 3520, Ap	62	US-10-425-115-3520	226	49.4	40	62
Sequence 63508, A	63	US-10-108-260A-3520	234	49.4	40	63
Sequence 1615, Ap	64	US-10-425-114-63508	491	49.4	40	64
Sequence 2149, Ap	65	US-10-408-765A-1615	713	49.4	40	65
Sequence 200705,	66	US-10-408-765A-2149	828	49.4	40	66
Sequence 226997,	67	US-10-425-115-200705	899	49.4	40	67
Sequence 9, Appli	68	US-10-424-599-226997	954	49.4	40	68
Sequence 122, App	69	US-10-205-342-9	1020	49.4	40	69
Sequence 5791, Ap	70	US-10-712-124-122	1020	49.4	40	70
Sequence 810, App	71	US-10-756-149-5791	1035	49.4	40	71
Sequence -807, App	72	US-10-322-281-810	1048	49.4	40	72
Sequence 169, App	73	US-10-322-281-807	2053	49.4	40	73
Sequence 11, Appl	74	US-10-051-874-169	4823	49.4	40	74
Sequence 8, Appli	75	US-10-142-515-11	5877	48.8	39.5	75
Sequence 5, Appli	76	US-10-243-243A-8	5935	48.8	39.5	76
Sequence 17, Appl	77	US-10-715-066-5	22152	48.8	39.5	77
Sequence 22, Appl	78	US-10-803-541-17	39	48.1	39	78
Sequence 30358, A	79	US-10-803-541-22	20	48.1	39	79
Sequence 278007,	80	US-10-029-386-30358	45	48.1	39	80
Sequence 276708,	81	US-10-425-115-278007	58	48.1	39	81
Sequence 363861,	82	US-10-424-599-276708	70	48.1	39	82
Sequence 240125,	83	US-10-425-115-363861	84	48.1	39	83
	84	US-10-424-599-240125	85	48.1	39	84

85 39 48.1 110 15 US-10-383-201-116 Sequence 116, App  
86 39 48.1 114 15 US-10-093-463-172 Sequence 172, App  
87 39 48.1 167 16 US-10-437-963-155180 Sequence 155180,  
88 39 48.1 177 16 US-10-425-115-300214 Sequence 300214,  
89 39 48.1 189 16 US-10-425-115-265787 Sequence 265787,  
90 39 48.1 281 15 US-10-375-150-12 Sequence 12, Appl  
91 39 48.1 291 16 US-10-425-115-198860 Sequence 198860,  
92 39 48.1 295 17 US-10-732-923-13904 Sequence 13904, A  
93 39 48.1 346 15 US-10-424-599-204524 Sequence 204524,  
94 39 48.1 385 15 US-10-264-049-3036 Sequence 3036, Ap  
95 39 48.1 402 8 US-08-957-425-29 Sequence 29, Appl  
96 39 48.1 402 8 US-08-260-675-21 Sequence 21, Appl  
97 39 48.1 402 10 US-09-952-318A-21 Sequence 21, Appl  
98 39 48.1 402 14 US-10-122-026-6 Sequence 6, Appl  
99 39 48.1 402 14 US-10-050-050-21 Sequence 21, Appl  
100 39 48.1 402 14 US-10-164-279-45 Sequence 45, Appl

## ALIGNMENTS

RESULT 1  
US-10-803-541-2  
; Sequence 2, Application US/10803541  
; Publication No. US2005090438A1  
; GENERAL INFORMATION:  
; APPLICANT: Brodsky, Gary  
; TITLE OF INVENTION: Product and Methods for Diagnosis and Therapy for Cardiac and  
; FILE REFERENCE: 2848-53  
; CURRENT APPLICATION NUMBER: US/10/803,541  
; CURRENT FILING DATE: 2004-03-17  
; PRIOR APPLICATION NUMBER: 60/456,642  
; PRIOR FILING DATE: 2003-03-18  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-803-541-2

Query Match 100.0%; Score 81; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
| | | | | | | | | | | | | | |  
Db 1 LLGNSSPRTQSPQNC 15

RESULT 2  
US-10-803-541-20  
; Sequence 20, Application US/10803541  
; Publication No. US2005090438A1  
; GENERAL INFORMATION:  
; APPLICANT: Brodsky, Gary  
; TITLE OF INVENTION: Product and Methods for Diagnosis and Therapy for Cardiac and  
; FILE REFERENCE: 2848-53  
; CURRENT APPLICATION NUMBER: US/10/803,541  
; CURRENT FILING DATE: 2004-03-17  
; PRIOR APPLICATION NUMBER: 60/456,642  
; PRIOR FILING DATE: 2003-03-18  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 20  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-803-541-20

Query Match 100.0%; Score 81; DB 17; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLGNSSPRTQSPQNC 15  
| | | | | | | | | | | | | | |  
Db 1 LLGNSSPRTQSPQNC 15

RESULT 3  
US-09-315-355-46  
; Sequence 46, Application US/09315355  
; Publication No. US20030157482A1  
; GENERAL INFORMATION:  
; APPLICANT: KEESSEE, SUSAN  
; APPLICANT: OHAR, ROBERT  
; APPLICANT: WU, YING-JYE  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
; DETECTION OF CERVICAL CANCER  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Testa, Hurwitz & Thibault  
; STREET: 125 High St.  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/315,355  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GREENHALGH, DUNCAN A  
; REGISTRATION NUMBER: 38,678  
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 515 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-315-355-46

Query Match 100.0%; Score 81; DB 10; Length 515;  
Best Local Similarity 100.0%; Pred. No. 0.00041;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
| | | | | | | | | | | | | | |  
Db 498 LLGNSSPRTQSPQNC 512

RESULT 4  
US-10-848-572-46  
; Sequence 46, Application US/10848572  
; Publication No. US20050164313A1  
; GENERAL INFORMATION:  
; APPLICANT: Keessee, Susan K.  
; APPLICANT: Ohar, Robert  
; APPLICANT: Wu, Ying-Jye  
; TITLE OF INVENTION: Methods and Compositions for the Detection of Cervical Cancer  
; FILE REFERENCE: MTP-023C1  
; CURRENT APPLICATION NUMBER: US/10/848,572  
; CURRENT FILING DATE: 2004-05-18  
; PRIOR APPLICATION NUMBER: US 08/989,045

; PRIOR FILING DATE: 1997-12-11  
; PRIOR APPLICATION NUMBER: US 08/705,660  
; PRIOR FILING DATE: 1996-08-30  
; PRIOR APPLICATION NUMBER: US 09/315,355  
; PRIOR FILING DATE: 1999-05-17  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46  
; LENGTH: 515  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-848-572-46

Query Match 100.0%; Score 81; DB 18; Length 515;  
Best Local Similarity 100.0%; Pred. No. 0.00041;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQNC 15  
Db 498 LLGNSPRTQSPQNC 512

RESULT 5  
US-10-408-765A-2079  
; Sequence 2079, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2079  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-2079

Query Match 100.0%; Score 81; DB 16; Length 664;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQNC 15  
Db 647 LLGNSPRTQSPQNC 661

RESULT 6  
US-10-473-127-1658  
; Sequence 1658, Application US/10473127  
; Publication No. US20040236091A1  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370

; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1658  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-1658

Query Match 100.0%; Score 81; DB 16; Length 664;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQNC 15  
Db 647 LLGNSPRTQSPQNC 661

RESULT 7  
US-10-943-400-2  
; Sequence 2, Application US/10943400  
; Publication No. US20050059071A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as  
; APPLICANT: represented by the Secretary of the Department of Health and  
; APPLICANT: Human Services  
; APPLICANT: Eriksson, B. Maria H.  
; APPLICANT: Collins, Francis S.  
; APPLICANT: Gordon, Leslie B.  
; APPLICANT: Brown, W. Ted  
; TITLE OF INVENTION: LMNA GENE AND ITS INVOLVEMENT IN HUTCHINSON-GILFORD PROGERIA  
; TITLE OF INVENTION: SYNDROME (HGPS) AND ARTERIOSCLEROSIS  
; FILE REFERENCE: 4239-66648-02  
; CURRENT APPLICATION NUMBER: US/10/943,400  
; CURRENT FILING DATE: 2004-09-17  
; PRIOR APPLICATION NUMBER: US 60/419,541  
; PRIOR FILING DATE: 2002-10-18  
; PRIOR APPLICATION NUMBER: PCT/US2003/0033058  
; PRIOR FILING DATE: 2002-10-18  
; PRIOR APPLICATION NUMBER: US 60/463,084  
; PRIOR FILING DATE: 2002-04-14  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 2  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-943-400-2

Query Match 100.0%; Score 81; DB 17; Length 664;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSPRTQSPQNC 15  
Db 647 LLGNSPRTQSPQNC 661

RESULT 8  
US-10-803-541-4  
; Sequence 4, Application US/10803541  
; Publication No. US20050090438A1  
; GENERAL INFORMATION:  
; APPLICANT: Brodey, Gary  
; TITLE OF INVENTION: Product and Methods for Diagnosis and Therapy for Cardiac and  
; TITLE OF INVENTION: Skeletal Muscle Disorders  
; FILE REFERENCE: 2848-53  
; CURRENT APPLICATION NUMBER: US/10/803,541  
; CURRENT FILING DATE: 2004-03-17

; PRIOR APPLICATION NUMBER: 60/456,642  
; PRIOR FILING DATE: 2003-03-18  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 4  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-803-541-4

Query Match 100.0%; Score 81; DB 17; Length 664;  
Best Local Similarity 100.0%; Pred. No. 0.00053;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
Db 647 LLGNSSPRTQSPQNC 661

RESULT 9  
US-10-803-541-16  
; Sequence 16, Application US/10803541  
; Publication No. US2005090438A1  
; GENERAL INFORMATION:  
; APPLICANT: Brodsky, Gary  
; TITLE OF INVENTION: Product and Methods for Diagnosis and Therapy for Cardiac and  
; FILE REFERENCE: 2848-53  
; CURRENT APPLICATION NUMBER: US/10/803,541  
; CURRENT FILING DATE: 2004-03-17  
; PRIOR APPLICATION NUMBER: 60/456,642  
; PRIOR FILING DATE: 2003-03-18  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 16  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-803-541-16

Query Match 85.2%; Score 69; DB 17; Length 15;  
Best Local Similarity 86.7%; Pred. No. 0.00082;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
Db 1 LLGNSSPRTQSPQNC 15

RESULT 10  
US-10-803-541-21  
; Sequence 21, Application US/10803541  
; Publication No. US2005090438A1  
; GENERAL INFORMATION:  
; APPLICANT: Brodsky, Gary  
; TITLE OF INVENTION: Product and Methods for Diagnosis and Therapy for Cardiac and  
; FILE REFERENCE: 2848-53  
; CURRENT APPLICATION NUMBER: US/10/803,541  
; CURRENT FILING DATE: 2004-03-17  
; PRIOR APPLICATION NUMBER: 60/456,642  
; PRIOR FILING DATE: 2003-03-18  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 21  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-803-541-21

Query Match 85.2%; Score 69; DB 17; Length 18;  
Best Local Similarity 86.7%; Pred. No. 0.00098;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
Db 1 LLGNSSPRTQSPQNC 15

RESULT 11  
US-10-803-541-9  
; Sequence 9, Application US/10803541  
; Publication No. US2005090438A1  
; GENERAL INFORMATION:  
; APPLICANT: Brodsky, Gary  
; TITLE OF INVENTION: Product and Methods for Diagnosis and Therapy for Cardiac and  
; FILE REFERENCE: 2848-53  
; CURRENT APPLICATION NUMBER: US/10/803,541  
; CURRENT FILING DATE: 2004-03-17  
; PRIOR APPLICATION NUMBER: 60/456,642  
; PRIOR FILING DATE: 2003-03-18  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 9  
; LENGTH: 665  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-803-541-9

Query Match 85.2%; Score 69; DB 17; Length 665;  
Best Local Similarity 86.7%; Pred. No. 0.038;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
Db 648 LLGNSSPRTQSPQNC 662

RESULT 12  
US-10-437-963-132511  
; Sequence 132511, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 132511  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_34473C.1.pap  
US-10-437-963-132511

Query Match 55.8%; Score 45; DB 16; Length 162;  
Best Local Similarity 60.0%; Pred. No. 48;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
Db 93 LLGNSSPRTQSPQNC 107

RESULT 13

```
US-10-425-115-198650
; Sequence 198650, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 198650
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_112748C.1.pep
US-10-425-115-198650

Query Match      54.3%; Score 44; DB 16; Length 122;
Best Local Similarity 72.7%; Pred. No. 52;
Matches      8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      2 LGNSSPRTQSP 12
Db      62 LGSSAPRTQPP 72
|||:|||||

RESULT 14
US-10-424-599-145179
; Sequence 145179, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 145179
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_102114C.1.pep
US-10-424-599-145179

Query Match      54.3%; Score 44; DB 15; Length 275;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches      8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy      2 LGNSSPRTQSPON 14
Db      4 LGNEQPGTQIPEN 16
|||:|||||

RESULT 15
US-10-094-749-2862
; Sequence 2862, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
```

```
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2862
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2862

Query Match      54.3%; Score 44; DB 15; Length 286;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches      8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      6 SPRTQSPQNC 15
Db      71 SPRTQSPGCG 80
|||||||

RESULT 16
US-10-437-963-117902
; Sequence 117902, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 117902
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21264C.1.pep
US-10-437-963-117902

Query Match      54.3%; Score 44; DB 16; Length 826;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches      8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 LLGNSSPRTQSP 12
Db      403 LLGNEGPRTSDP 414
|||||||
```

```
RESULT 17
US-10-437-963-202808
; Sequence 202808, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 202808
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_98052C.1.pep
US-10-437-963-202808

Query Match      53.1%; Score 43; DB 16; Length 539;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches      8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1  LLGNSSPRTQSP 12
      :||| |||||
DB      409  LVNGSPRTTPP 420

RESULT 18
US-10-481-032A-392
; Sequence 392, Application US/10481032A
; Publication No. US20050177901A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Cheng, Wenqiong
; APPLICANT: Briggs, Steven
; APPLICANT: Cooper, Bret
; APPLICANT: Goff, Stephen A.
; APPLICANT: Moughamer, Todd
; APPLICANT: Glazebrook, Jane
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicolas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 60148USPCT
; CURRENT APPLICATION NUMBER: US/10/481,032A
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: US 60/300,112
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/342,327
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: PCT/IB02/02450
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 1201
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 392
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-481-032A-392
```

```
Query Match      53.1%; Score 43; DB 18; Length 558;
Best Local Similarity 57.1%; Pred. No. 3.5e+02;
Matches      8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1  LLGNSSPRTQSPQN 14
      :||| |||||
DB      231  ILQSNIPRTQPPQN 244

RESULT 19
US-10-437-963-203613
; Sequence 203613, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203613
; LENGTH: 1797
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1797)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_98780C.1.pep
US-10-437-963-203613

Query Match      53.1%; Score 43; DB 16; Length 1797;
Best Local Similarity 57.1%; Pred. No. 1.1e+03;
Matches      8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1  LLGNSSPRTQSPQN 14
      :||| |||||
DB      210  ILQSNIPRTQPPQN 223

RESULT 20
US-10-437-963-203615
; Sequence 203615, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203615
; LENGTH: 2001
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
```

OTHER INFORMATION: Clone ID: PAT\_MRT4530\_98782C.1.pep  
US-10-437-963-203615

Query Match 53.1%; Score 43; DB 16; Length 2001;  
Best Local Similarity 57.1%; Pred. No. 1.3e+03;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQN 14  
Db 231 ILQSNIPRTQPPQN 244

## RESULT 21

US-10-767-701-46748  
; Sequence 46748, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 46748  
; LENGTH: 202  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C49\_82.pep  
US-10-767-701-46748

Query Match 52.5%; Score 42.5; DB 16; Length 202;  
Best Local Similarity 69.2%; Pred. No. 1.5e+02;  
Matches 9; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 3 GNSPRTQSPQNC 15  
Db 154 GNISPLTQ-PENC 165

## RESULT 22

US-10-424-599-234564  
; Sequence 234564, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 234564  
; LENGTH: 81  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_53839C.1.pep  
US-10-424-599-234564

Query Match 51.9%; Score 42; DB 15; Length 81;  
Best Local Similarity 46.7%; Pred. No. 70;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQNC 15  
Db 24 LCGAPPRTSPRC 38

## RESULT 23

US-10-425-115-324347  
; Sequence 324347, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 324347  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_58879C.1.pep  
US-10-425-115-324347

Query Match 51.9%; Score 42; DB 16; Length 100;  
Best Local Similarity 61.5%; Pred. No. 87;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LLGNSSPRTQSPQ 13  
Db 54 LVGNSSPNLGTGPQ 66

## RESULT 24

US-10-283-940-20  
; Sequence 20, Application US/10283940  
; Publication No. US20030220394A1  
; GENERAL INFORMATION:  
; APPLICANT: Danisco A/S  
; APPLICANT: Morgan, Andrew J  
; APPLICANT: Yu, Shukun  
; APPLICANT: Weiergang, Inge  
; APPLICANT: Pedersen, Hans C  
; TITLE OF INVENTION: Sequences  
; FILE REFERENCE: 674509-2041.1  
; CURRENT APPLICATION NUMBER: US/10/283,940  
; CURRENT FILING DATE: 2002-10-30  
; PRIOR APPLICATION NUMBER: PCT/GB02/04916  
; PRIOR FILING DATE: 2002-10-30  
; PRIOR APPLICATION NUMBER: GB 0126164.3  
; PRIOR FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US 60/343,485  
; PRIOR FILING DATE: 2001-12-21  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 175  
; TYPE: PRT  
; ORGANISM: Phanerochaete chrysosporium  
US-10-283-940-20

Query Match 51.9%; Score 42; DB 15; Length 175;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LGNSSPRTQSPQNC 15  
Db 67 IGLGSPITETQRC 80

## RESULT 25

US-11-022-454-20

```

; Sequence 20, Application US/11022454
; Publication No. US20050164259A1
; GENERAL INFORMATION:
; APPLICANT: Morgan, Andrew John Pedersen, Hans Christian
; APPLICANT: Morgan, Andrew John
; APPLICANT: Weiergang, Inge
; APPLICANT: Yu, Shukun
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 674509-2041.2
; CURRENT APPLICATION NUMBER: US/11/022,454
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 10/283,940
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: PCT/GB02/04916
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: GB 0126164.3
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 60/343,485
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Phanerochaete chrysosporium
US-11-022-454-20

Query Match 51.9%; Score 42; DB 20; Length 175;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 LGNSSPRTQSPQNC 15
      : ||| :|||
Db 67 IGLGPITETPQC 80

RESULT 26
US-09-994-064-19
; Sequence 19, Application US/09994064
; Publication No. US20030082788A1
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUSLYRNAGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/994,064
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,190
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:

```

; Sequence 19, Application US/10836383  
; Publication No. US20040228878A1  
; GENERAL INFORMATION:  
; APPLICANT: Wild, Martha A.  
; COCHRAN, Mark D.  
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS  
; AND USES THEREOF  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/836,383  
; FILING DATE: 30-Apr-2004  
; CLASSIFICATION DATA:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,190C  
; FILING DATE: 06-JUN-1995  
; APPLICATION NUMBER: US 08/126,597  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 39116-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 177 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-10-836-383-19  
Query Match 51.9%; Score 42; DB 16; Length 177;  
Best Local Similarity 53.3%; Pred. No. 1.5e+02;  
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 LLGNSPRTQSPQNC 15  
DB 5 LLNRGSPRLNSPPKC 19  
RESULT 29  
US-10-424-599-206208  
; Sequence 206208, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 206208  
; LENGTH: 183  
; TYPE: PRT  
; ORGANISM: Glycine max

; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_28232C.1.pep  
US-10-424-599-206208  
Query Match 51.9%; Score 42; DB 15; Length 183;  
Best Local Similarity 72.7%; Pred. No. 1.6e+02;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 4 NSSPRTQSPQN 14  
DB 51 NSNPQAQSPQN 61  
RESULT 30  
US-10-104-047-3255  
; Sequence 3255, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20030236392A1e1 full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3255  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-3255  
Query Match 51.9%; Score 42; DB 15; Length 256;  
Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 3 GNSPRTQSP 12  
DB 234 GNSPHTASP 243  
RESULT 31  
US-10-437-963-203609  
; Sequence 203609, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 203609  
; LENGTH: 1701  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_98777C.1.pep  
US-10-437-963-203609  
Query Match 51.9%; Score 42; DB 16; Length 1701;  
Best Local Similarity 57.1%; Pred. No. 1.5e+03;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
QY      1 LLGNSSPRTQSPQN 14
      :|::|||
Db      210 ILQSNIPRTQLPQN 223

RESULT 32
US-10-437-963-204611
; Sequence 204611, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204611
; LENGTH: 1828
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1) .. (1828)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99681C.1.pep
US-10-437-963-204611

Query Match      51.9%; Score 42; DB 16; Length 1828;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;

QY      1 LLGNSSPRTQSPQN 14
      :|::|||
Db      210 ILQSNIPRTQLPQN 223

RESULT 33
US-10-437-963-203618
; Sequence 203618, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203618
; LENGTH: 1980
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1) .. (1980)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
```

```
; OTHER INFORMATION: Clone ID: PAT_MRT4530_98785C.1.pep
US-10-437-963-203618

Query Match      51.9%; Score 42; DB 16; Length 1980;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;
Matches      8; Conservative      3; Mismatches      3; Indels      0; Gaps      0;

QY      1 LLGNSSPRTQSPQN 14
      :|::|||
Db      231 ILQSNIPRTQLPQN 244

RESULT 34
US-10-135-322-17
; Sequence 17, Application US/10135322
; Publication No. US20020173017A1
; GENERAL INFORMATION:
; APPLICANT: BENFEY, PN
; APPLICANT: HELARIUTTA, Y
; APPLICANT: MAHONEN, AP
; APPLICANT: BONKE, AMW
; APPLICANT: KAUPPINEN, L
; APPLICANT: RIIKONEN, M
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF
; FILE REFERENCE: 5914-086-999
; CURRENT APPLICATION NUMBER: US/10/135,322
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/253,739
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-135-322-17

Query Match      51.9%; Score 42; DB 13; Length 2150;
Best Local Similarity 53.8%; Pred. No. 2e+03;
Matches      7; Conservative      2; Mismatches      4; Indels      0; Gaps      0;

QY      3 GNSSPRTQSPQNC 15
      ||::||
Db      345 GNSSPRNCGSNC 357

RESULT 35
US-10-425-115-218829
; Sequence 218829, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 218829
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_131163C.1.pep
US-10-425-115-218829

Query Match      50.6%; Score 41; DB 16; Length 78;
Best Local Similarity 70.0%; Pred. No. 96;
Matches      7; Conservative      1; Mismatches      2; Indels      0; Gaps      0;
```

```
Qy      6 SPRTSQPNC 15
Db      22 SPMTQSPRC 31

RESULT 36
US-10-425-115-252593
; Sequence 252593, Application US/10425115
; Publication No. US2004021472A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 252593
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_161948C.1.pep
US-10-425-115-252593

Query Match      50.6%; Score 41; DB 16; Length 86;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy      3 GNSSPR--TQSPNC 15
Db      15 GVSQPRVCTQAPMNC 29

RESULT 37
US-10-424-599-230506
; Sequence 230506, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 230506
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_50169C.1.pep
US-10-424-599-230506

Query Match      50.6%; Score 41; DB 15; Length 112;
Best Local Similarity 63.6%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      5 SSPRTQSPNC 15
Db      55 STPRSSDDQNC 65

RESULT 38
US-10-767-701-57287
; Sequence 57287, Application US/10767701
; Publication No. US20040172684A1

; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 57287
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30965256.pep
US-10-767-701-57287

Query Match      50.6%; Score 41; DB 16; Length 112;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy      3 GNSSPRTQSPON 14
Db      1 GSRSPRSRSPQD 12

RESULT 39
US-10-425-114-38640
; Sequence 38640, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38640
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700347220_FLI.pep
US-10-425-114-38640

Query Match      50.6%; Score 41; DB 15; Length 243;
Best Local Similarity 58.3%; Pred. No. 3.1e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy      3 GNSSPRTQSPON 14
Db      118 GSRSPRSRSPQD 129

RESULT 40
US-10-425-114-47848
; Sequence 47848, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 47848  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700267149\_FLI.pep  
US-10-425-114-47848

Query Match 50.6%; Score 41; DB 15; Length 310;  
Best Local Similarity 58.3%; Pred. No. 3.9e+02;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPQN 14  
|:|:|:|:|:  
Db 199 GSRSPRSRSPQD 210

## RESULT 41

US-10-425-114-60855  
; Sequence 60855, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 60855  
; LENGTH: 316  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3354-031-A12\_FLI.pep  
US-10-425-114-60855

Query Match 50.6%; Score 41; DB 15; Length 316;  
Best Local Similarity 58.3%; Pred. No. 4e+02;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPQN 14  
|:|:|:|:|:  
Db 211 GSRSPRSRSPQD 222

## RESULT 42

US-10-425-114-69314  
; Sequence 69314, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 69314  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17123H02\_FLI.pep  
US-10-425-114-69314

Query Match 50.6%; Score 41; DB 15; Length 320;  
Best Local Similarity 58.3%; Pred. No. 4e+02;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPQN 14  
|:|:|:|:|:  
Db 195 GSRSPRSRSPQD 206

## RESULT 43

US-10-425-115-257256  
; Sequence 257256, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 257256  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_166202C.1.pep  
US-10-425-115-257256

Query Match 50.6%; Score 41; DB 16; Length 328;  
Best Local Similarity 58.3%; Pred. No. 4.1e+02;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GNSSPRTQSPQN 14  
|:|:|:|:|:  
Db 203 GSRSPRSRSPQD 214

## RESULT 44

US-10-425-115-258656  
; Sequence 258656, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 258656  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_167481C.1.pep  
US-10-425-115-258656

Query Match 50.6%; Score 41; DB 16; Length 334;

Best Local Similarity 58.3%; Pred. No. 4.2e+02;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 GNSSPRTQSPQN 14  
| : | | | : | | :  
Db 228 GSRSPRSRSPQD 239

```

RESULT 45
US-10-425-115-258661
; Sequence 258661, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 258661
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(338)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167486C.1.dep
; US-10-425-115-258661

```

```

RESULT 46
US-10-425-115-257255
; Sequence 257255, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION: Thomas J.
; APPLICANT: La Rosa,
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OP INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 257255
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Cione ID: MPT4577_166201C.1.pgp
; US-10-425-115-257255

```

Db 228 GSRSPRSRPQD 239

```

RESULT 47
US-10-425-114-58835
; Sequence 58835, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58835
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700221865_FLI.pep
US-10-425-114-58835

```

```

RESULT 48
US-10-425-115-258659
; Sequence 258659, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 258659
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_l67484C.1.pep
US-10-425-115-258659

```

RESULT 49  
US-10-425-114-60636  
; Sequence 60636, Application US/10425114  
; Publication No. US20040034888A1

```
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60636
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-001-A11_FLI.pep
US-10-425-114-60636

Query Match          50.6%; Score 41; DB 15; Length 365;
Best Local Similarity 58.3%; Pred. No. 4.6e+02;
Matches              7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      3   GNSSPRTQSPOK 14
       |:|::|||:
Db      262  GSRSPRSRSPQD 273

RESULT 50
US-10-425-114-59471
; Sequence 59471, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59471
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3151-024-C2_FLI.pep
US-10-425-114-59471

Query Match          50.6%; Score 41; DB 15; Length 370;
Best Local Similarity 58.3%; Pred. No. 4.7e+02;
Matches              7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      3   GNSSPRTQSPOK 14
       |:|::|||:
Db      255  GSRSPRSRSPQD 266

Search completed: August 25, 2005, 11:48:10
Job time : 163 secs
```